

**Entwicklung eines
impulskodierenden neuronalen Netzes
für die Segmentierung bewegter Szenen**

Dissertation
zur
Erlangung des Doktorgrades
der Naturwissenschaften
(Dr. rer. nat.)

dem
Fachbereich Physik
der Philipps-Universität Marburg
vorgelegt von

Ulrich Schott

aus Münchenhausen

Marburg/Lahn 1999

Vom Fachbereich Physik der Philipps-Universität Marburg als Dissertation am 4. Mai 1999

angenommen

Erstgutachter: Prof. Dr. Dr. H. J. Reitböck

Zweitgutachter: Prof. Dr. R. Eckhorn

Tag der mündlichen Prüfung: 5. Juli 1999

Inhaltsverzeichnis

1 Einleitung	1
2 Grundlagen	5
2.1 Psychophysische Grundlagen	5
2.1.1 Apertur- und Korrespondenzproblem	5
2.1.2 Was erzeugt eine eindeutige Bewegungswahrnehmung?	6
2.1.2.1 Korrespondenz durch Objekteigenschaften	6
2.1.2.2 Korrespondenz durch zeitliche Integration der Bewegung	7
2.1.2.3 Korrespondenz durch räumliche Integration der Bewegung	9
2.1.2.4 Einwirkung der Terminatoren auf die Bewegungswahrnehmung	10
2.1.3 Szenensegmentierung durch Bewegungsinformation	13
2.2 Biologische Grundlagen	14
2.2.1 Das biologische Neuron	14
2.2.2 Die Retina	16
2.2.3 Die visuellen Verarbeitungsbereiche des Kortex	18
2.3 Grundlagen zur Modellierung	23
2.3.1 Methoden der Bewegungsdetektion	23
2.3.2 Das Marburger Modellneuron	27
2.3.3 Der Hardware-Accelerator	29
2.3.3.1 Motivation und Simulationsumgebung	29
2.3.3.2 Zeitschritt-Simulationsverfahren	31
2.3.3.3 Module des Hardware-Accelerators	32
2.3.3.4 Netztopologie und Gewichtespeicher	33
2.3.3.5 Die Berechnung des Membranpotentials	34
3 Modelle und Simulationsergebnisse	37
3.1 Die retinale Vorverarbeitung	37
3.1.1 Das Vorverarbeitungsmodell	37
3.1.1.1 Zeitliche Interpolation der Kamerabilder	37
3.1.1.2 Hexagonale Unterabtastung der Eingabebilder	38
3.1.1.3 Das Rezeptorpotential	38
3.1.1.4 Das Bipolarzellenmodell	40
3.1.1.5 Das Ganglienzellmodell	44
3.1.2 Simulationsergebnisse der Vorverarbeitung	48
3.1.2.1 Simulationsergebnisse mit künstlichen Szenen	48
3.1.2.2 Simulationsergebnisse mit realen Szenen	64
3.2 Bewegungsdetektion mit impulskodierenden neuronalen Netzen	70
3.2.1 Modell zur Bewegungsdetektion	70
3.2.1.1 Modell des Richtungsdetektors	72
3.2.1.2 Laterale, unverzögerte Kopplung der Richtungsdetektoren	74
3.2.1.3 Laterale, verzögerte Kopplung der Richtungsdetektoren	78
3.2.1.4 Kopplung von ON- und OFF-Pfad	80

3.2.2 Simulationsergebnisse der Bewegungsdetektion.....	82
3.2.2.1 Simulationsergebnisse mit künstlichen Szenen.....	82
3.2.2.2 Simulationsergebnisse mit realen Szenen.....	102
3.3 Kopplung von Richtungsdetektoren unterschiedlicher Auflösung.....	119
3.3.1 Modell zur Kopplung von Richtungsdetektoren unterschiedlicher Auflösung...	119
3.3.2 Simulationen zur Auflösungskopplung.....	123
3.3.2.1 Simulationsergebnisse mit künstlichen Szenen.....	123
3.3.2.2 Simulationsergebnisse mit realen Szenen.....	129
3.4 Einbindung der Linienenden in die Bewegungsdetektion.....	133
3.4.1 Gesamtmodell zur Einbindung der Linienenden in die Bewegungsverarbeitung	133
3.4.1.1 Modell zur Orientierungsdetektion.....	136
3.4.1.2 Modell zur Linienendendetektion.....	143
3.4.1.3 Modell zur Bestimmung der Linienendenbewegungsrichtung.....	147
3.4.2 Simulationen zur Einbindung der Linienenden in die Bewegungsverarbeitung..	154
3.4.2.1 Simulation der Barberpole Illusion und der Barberplaid Illusion.....	154
3.4.2.2 Linienendendetektion in realen Szenen.....	168
3.5 Bewegungskontrast.....	172
3.5.1 Modell zur Extraktion des lokalen Bewegungskontrastes.....	172
3.5.2 Simulationen zur Bewegungskontrastextraktion.....	177
4 Diskussion.....	183
4.1 Vorverarbeitung.....	183
4.2 Detektion der Konturbewegungsrichtung.....	184
4.2.1 Einordnung in psychophysische Untersuchungen.....	184
4.2.2 Einordnung in biologische Untersuchungen.....	184
4.2.3 Vergleich mit anderen Modellen.....	186
4.2.4 Motion Capture durch Richtungsdetektoren unterschiedlicher Auflösung.....	187
4.2.5 Objektverfolgung.....	188
4.3 Barberpole und Barberplaid Illusion.....	188
4.3.1 Einordnung in psychophysische Untersuchungen.....	188
4.3.2 Einordnung in biologische Untersuchungen.....	189
4.3.3 Vergleich mit anderen Modellen.....	190
4.4 Bewegungskontrast.....	191
4.5 Einsetzbarkeit in der technischen Bildverarbeitung.....	192
4.6 Ausblick.....	193
5 Literatur.....	195

Zusammenfassung

Die Detektion lokaler Bewegungsrichtungen in visuellen Szenen liefert aufgrund des Apertur- und des Korrespondenzproblems mehrdeutige Ergebnisse. Biologische Sehsysteme haben aber eindeutige Wahrnehmungen und können anhand der Bewegungsinformation eine Objektseparation durchführen. Wahrnehmungs- und neurowissenschaftliche Untersuchungen geben Hinweise, wie biologische Sehsysteme diese komplexen Aufgaben lösen könnten.

In dieser Arbeit wird ein Netz mit biologienahen Modellneuronen zur Bewegungsrichtungsdetektion entwickelt, in dessen Verschaltung und Aktivität sich psychophysische und neurophysiologische Erkenntnisse, insbesondere zur zeitlichen und räumlichen Integration von Bewegungsinformation, widerspiegeln. Die Funktionalität des neuronalen Netzes entspricht der des bewegungssensitiven Verarbeitungspfades im Sehsystem von Primaten.

Das hierarchisch aufgebaute neuronale Netz detektiert Merkmalskontraste. Es werden in zwei parallelen Verarbeitungspfaden sowohl die Bewegungsrichtung von Objektkonturen als auch die Bewegungsrichtung von scharfen Konturkrümmungen extrahiert. An den scharfen Krümmungen oder Ecken der Kontur kann im Gegensatz zu den geraden Konturstücken eine eindeutige Bewegungsrichtung bestimmt werden. Dazu werden lokale Intensitätsgradienten und zeitliche Transienten entsprechend einer retinalen Vorverarbeitung herausgefiltert. Anschließend wird im ersten Pfad mit bilokalen Richtungsdetektoren die Konturbewegungsrichtung extrahiert. Nichtlinear wirkende laterale Koppel-Verbindungen zwischen den Richtungsdetektoren und Rückkopplungen über inhibitorisch wirkende Interneuronen unterstützen dabei die Erzeugung einer eindeutigen Richtungsdetektion, indem sie eine zeitliche und räumliche Integration der Bewegungsinformation ermöglichen.

Im zweiten Pfad werden aus orientierten Kantensegmenten Linienenden und deren Bewegungsrichtung bestimmt. Mit der Bewegungsrichtung der Linienenden wird nachfolgend über laterale Kopplung und einen Wettstreit über rückgekoppelte Hemmung die Bewegungsrichtungsdetektion der Kontur so beeinflusst, daß die Konturbewegung und die Linienendenbewegung zu einer einheitlichen Aussage für zusammengehörige Objektmerkmale führen. Mit dieser neuronalen Verschaltung können Wahrnehmungseffekte wie die „barberpole illusion“ und die „barberplaid illusion“ erstmals explizit durch eine neuronale Verschaltungsstruktur erklärt werden.

Aus der detektierten Bewegungsrichtung wird in nachfolgenden Netzwerkstufen der lokale Bewegungskontrast bestimmt. Der Bewegungskontrast kann zur Objektseparation vom Hintergrund herangezogen werden.

Mit dem in dieser Arbeit entwickelten neuronalen Netz können nicht nur psychophysische Effekte erklärt und Vorhersagen über eine mögliche Verschaltungsstruktur in biologischen Sehsystemen gemacht werden, sondern es ist auch in der technischen Bildverarbeitung erfolgreich einsetzbar, wie an zahlreichen Simulationen mit realen Szenen des Straßenverkehrs gezeigt werden kann.

1 Einleitung

Computer sind in den letzten Jahrzehnten immer mehr Teil unseres Lebens geworden. Sie unterstützen uns sowohl beruflich als auch privat, eröffnen uns ungeahnte Möglichkeiten und sind aus unserem Alltag nicht mehr wegzudenken. Doch fast alle Computer haben einen gemeinsamen Nachteil: Die Kommunikation mit ihnen ist beschränkt, da sie weder sehen, hören noch sprechen können. Computer, die diese Fähigkeiten rudimentär besitzen, sind auf eine ganz spezielle Aufgabe restringiert. Die Behebung dieses gravierenden Nachteils, insbesondere der fehlenden visuellen Fähigkeiten der Computer, ist seit vielen Jahren Gegenstand der Forschung. Erfolge wurden aber bei der visuellen Informationsverarbeitung mit technischen Systemen bisher nur mit einzelnen, problemspezifischen Anwendungen erzielt. Ein universell einsetzbares technisches System gibt es bislang nicht. Die Natur zeigt uns aber, daß universell einsetzbare Systeme zur visuellen Informationsverarbeitung existieren: Ein Beispiel ist das visuelle System des Menschen. Es liegt somit nahe, die Erkenntnisse, die über biologische Sehsysteme vorliegen, in die Konstruktion eines technischen Systems zur visuellen Informationsverarbeitung einfließen zu lassen. Die Untersuchung biologischer Sehsysteme und das Erstellen geeigneter Modelle ist aber wegen der großen Anzahl an Verarbeitungseinheiten und Verbindungen, die zu einer immensen Anzahl an möglichen Verschaltungen führen, schwierig. Einblicke in die Funktionsweise des Gehirns bieten hierbei insbesondere neurophysiologische, aber auch anatomische und psychophysische Untersuchungen. Diese geben wichtige Hinweise zur Gestaltung der Modelle, zu denen vor allem auch die technischen neuronalen Netze gehören.

Ein grundlegendes Problem der visuellen Informationsverarbeitung ist die Generierung einer invarianten Objektrepräsentation. Aus einer Fülle von Informationen müssen die für das Objekt charakteristischen Merkmale extrahiert werden. Allerdings stellt sich hier bereits die Frage, welche Bereiche und Merkmale in einer Szene zu welchem Objekt gehören. Einerseits kann ein Objekt erst erkannt werden, wenn es von den anderen Objekten und vom Hintergrund separiert worden ist, andererseits kann das Objekt erst vom Hintergrund und von anderen Objekten separiert werden, wenn die Merkmale eines Objektes zueinander gruppiert wurden und das Objekt erkannt wurde. An dieses Dilemma schließt sich gleich eine weitere Frage an, wie nämlich die Zusammengehörigkeit der Merkmale eines Objektes in der internen Repräsentation bestimmt und kodiert wird; es ist dies das sogenannte „*Binding-Problem*“. Die Entwicklung eines universell einsetzbaren technischen Systems erfordert die Lösung dieser Probleme.

Gestaltpsychologen fanden schon in den zwanziger und dreißiger Jahren erste Hinweise zur Lösung des Dilemmas der Objektseparation. Die Gruppierung von Merkmalen und Objekten scheint nach bestimmten Prinzipien zu erfolgen. Diese sind das Prinzip der *Nähe*, des *gemeinsamen Schicksals*, des *guten Verlaufs*, der *ähnlichen Größe*, der *ähnlichen Form* und der *Geschlossenheit*. Weiterhin können Gruppierung und Trennung aufgrund der *Farbe*, der *Intensität*, der *Textur*, der *Disparität* und der *Bewegung* erfolgen. In diesen Domänen nehmen die lokalen Kontraste - das sind die Bereiche hohen Informationsgehaltes - sowohl für die Objektseparation als auch für die Merkmalsextraktion eine besondere Stellung ein.

Einen möglichen Ansatz zur Klärung des Binding-Problems liefert die *Korrelationshypothese*. Diese zu Beginn der achtziger Jahre erstellte Hypothese führt die Zusammengehörigkeit in der internen Repräsentation auf eine synchrone neuronale Aktivität zurück (Reitboeck 1980, Reitboeck 1983, von der Malsburg 1981, von der Malsburg und Schneider 1986). Die Hypothese wird zum einen durch die Entdeckung neuronaler Synchronisationsvorgänge im visuellen System der Katze (Eckhorn et al. 1988, Gray et al. 1989, Singer und Gray 1995), zum anderen durch die erfolgreiche Anwendung des Prinzips in der technischen Szenensegmentierung (Reitboeck et al. 1993, Reitboeck und Stoecker 1993, Stoecker 1993, Schillen und König 1991, von der Malsburg und Buhmann 1992, Stoecker und Reitboeck 1993, Stoecker et al. 1996, Weitzel et al. 1997a, Weitzel et al. 1997b, Weitzel 1998) unterstützt.

Das in dieser Arbeit entwickelte neuronale Netz ist Teil eines an biologischen Sehsystemen orientierten technischen Systems, das auf Ergebnissen neurophysiologischer und psychophysischer Untersuchungen basiert und das in der technischen Bildverarbeitung eingesetzt werden kann. Das in Abbildung 1.1 dargestellte Blockschaltbild führt unter Verwendung der Gestaltprinzipien eine Objektseparation im Sinne der Korrelationshypothese durch. Das neuronale Netz ist hierzu in vier Module aufgeteilt: Vorverarbeitung, Kontur/Form-System, Transientensystem und Aufmerksamkeitssteuerung. Das Kontur/Form-System (Weitzel et al. 1996, Weitzel et al. 1997a, Weitzel et al. 1997b, Weitzel 1998) hat dabei die Aufgabe, eine auf extrahierten Merkmalen basierende Objektdefinition durch synchrone neuronale Aktivität durchzuführen. Um dies erfolgreich bewerkstelligen zu können, benötigt es hochauflösende Verarbeitungseinheiten und ausreichend Verarbeitungszeit. Da nur eine begrenzte Anzahl an Verarbeitungseinheiten zur Verfügung steht, kann nur ein kleiner Bereich des Blickfeldes verarbeitet werden. Dieser hochauflösende Verarbeitungsbereich muß nun auf mögliche Objektkandidaten im Blickfeld gesetzt werden und sie gegebenenfalls zur genaueren Analyse verfolgen. Dazu dienen das in dieser Arbeit entwickelte Transientensystem (Fenske et al. 1995, Schott et al. 1996, Schott et al. 1997, Schott und Eckhorn 1998) und eine Blicksteuerung (Pauly et al. 1997, Pauly et al. 1998). Die Vorverarbeitung setzt die von der Kamera aufgenommenen Eingabebilder in neuronale Aktivitäten um und führt dabei eine erste Datenreduktion durch.

Das Transientensystem ist wie das Kontur/Form-System hierarchisch mit *Marburger Modellneuronen* (Eckhorn et al. 1988) aufgebaut und hat die Aufgabe, Bewegungsinformation aus den Eingabebildern zu extrahieren. Dies erfolgt über eine bilokale Bewegungsdetektion (Reichardt 1961). Die Bewegungsdetektion unterliegt aber grundlegenden Problemen: dem Apertur- und dem Korrespondenzproblem (Wallach 1935). Sie beinhalten das Fehlen der für eine eindeutige Bewegungsdetektion erforderliche Korrespondenz. Dennoch ist die menschliche Bewegungswahrnehmung im allgemeinen eindeutig. Psychophysische Experimente geben auch hier Hinweise zur Problemlösung. Diese Arbeit zeigt mögliche neuronale Verschaltungen auf, die für viele Situationen zu einem eindeutigen Ergebnis der Bewegungsdetektion führen könnten. In diesem Zusammenhang sind besonders die *barberpole illusion* und die *barberplaid illusion* von Bedeutung. An ihr läßt sich erkennen, daß markante Stellen des sich bewegenden Objektes, wie zum Beispiel Ecken, wichtig für die Bewegungswahrnehmung sind, denn diese unterliegen nur im geringen Maße dem Korrespondenzproblem, wodurch an diesen Stellen die Bewegungsrichtung des Objektes eindeutig bestimmt werden kann. Neben der Bestimmung der Bewegungsrichtung der Kontur ist somit die Extraktion der Linienenden und deren Bewegungsrichtung ein wesentlicher Bestandteil dieser Arbeit.

Das ganze in Abbildung 1.1 dargestellte System ist mit Marburger Modellneuronen aufgebaut. Da dieses technische System der Echtzeitanforderung möglichst nahekommen soll, wurde eine spezielle Hardware zur beschleunigten Berechnung neuronaler Netze („Accelerator“) entwickelt (Frank und Hartmann 1995, Frank et al. 1995, Frank 1997). Alle Module mit Ausnahme der Vorverarbeitung sind konform zur Hardware implementiert und unterliegen somit den Beschränkungen bezüglich der Modellneuronen- und der Verbindungsanzahl sowie der Rechengenauigkeit.

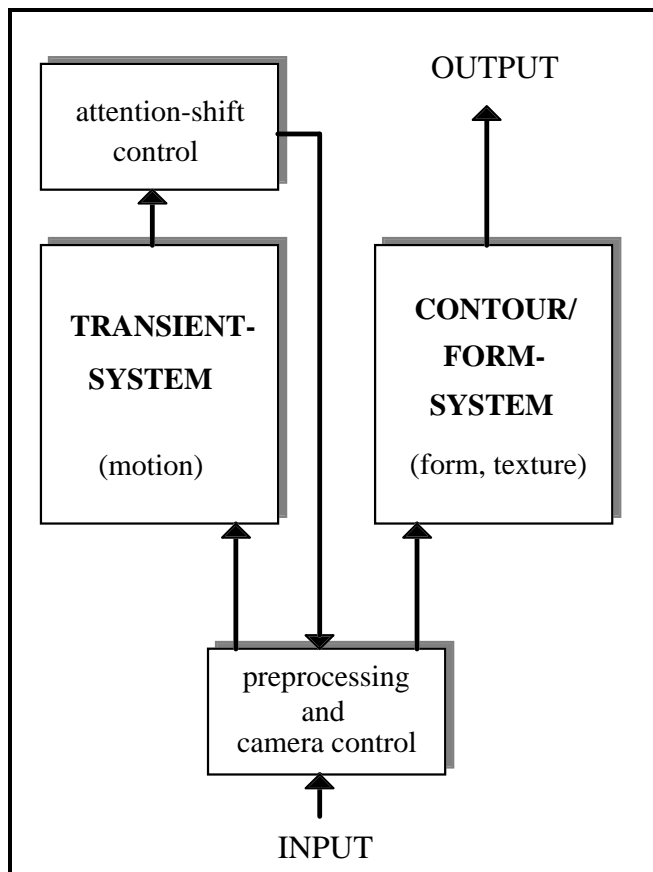


Abb. 1.1. Blockschaltbild zur Objekt/Hintergrund-Separation.

Dargestellt sind die vier Module des Gesamtsystems. Das Kontur/Form-System generiert eine interne, auf Merkmalen basierende Objektrepräsentation durch synchrone neuronale Aktivität. Das Transientensystem extrahiert aus den Eingabebildern die Bewegungsinformation, wodurch es ebenfalls zur Objektseparation und zur Aufmerksamkeitssteuerung beiträgt. Die Vorverarbeitung sowie die Ansteuerung der Kamera werden mit Hilfe eines konventionellen Computers durchgeführt. Die drei anderen Verarbeitungsmodulen sind auf einer speziellen Hardware (Accelerator) implementiert.

In Kapitel 2 dieser Arbeit werden zunächst die psychophysischen (Kap. 2.1) und neurophysiologischen (Kap. 2.2) Grundlagen zur Bewegungsverarbeitung dargestellt. Dies umfaßt sowohl die in der Bewegungsverarbeitung auftretenden Probleme als auch Untersuchungen, die zur Lösung dieser Probleme beitragen könnten. Weiterhin werden in Kapitel 2.3 grundlegende Modelle zur Bewegungsverarbeitung beschrieben. Prinzipielle Eigenschaften dieser Modelle finden sich in dem in dieser Arbeit entwickelten Modell zur Bewegungsverarbeitung mit impulskodierenden Modellneuronen wieder. In Kapitel 2.3.3 wird zusätzlich ein Überblick über das Acceleratorsystem und dessen Einschränkungen bezüglich der Simulationen gegeben.

Das dritte Kapitel umfaßt sowohl die Beschreibung des in dieser Arbeit entwickelten neuronalen Netzwerks als auch die Ergebnisse der Verarbeitung realer Szenen und künstlicher

Stimuli. Der Aufbau des gesamten Modells wird mit Simulationsergebnissen in den Unterkapiteln sukzessive dargestellt.

Zunächst wird in Kapitel 3.1 die Vorverarbeitung beschrieben. Die Vorverarbeitung stellt das Bindeglied zwischen Kamera und Acceleratorsystem dar und entspricht in ihrer Funktion in vielen Eigenschaften der der Retina. Das neuronale Netz zur Detektion der Bewegungsrichtung wird in Kapitel 3.2 vorgestellt. Dieses für die weitere Arbeit grundlegende Modell erzeugt mittels lateraler Kopplungen zwischen den Modellneuronen ein eindeutiges Ergebnis bei der Bewegungsrichtungsdetektion für Konturen. Mit diesem neuronalen Netz können bereits einige psychophysische Effekte erklärt werden. Die Kopplung von Modellneuronen mit unterschiedlicher räumlicher Auflösung wird in Kapitel 3.3 untersucht. In Kapitel 3.4 wird das neuronale Netz weiterentwickelt, so daß sich die Bewegungsrichtung der Linienenden eindeutig extrahieren läßt. Die Bewegungsrichtung der Linienenden wird über laterale Kopplungen auf die zugehörigen Konturen übertragen, an der sie zu einem eindeutigen Ergebnis der Bewegungsrichtungsdetektion, das der der Linienenden entspricht, beiträgt. Mit dieser Verschaltung können sowohl die *barberpole illusion* als auch die *barberplaid illusion* erklärt werden. In Kapitel 3.5 wird schließlich ein neuronales Netz entwickelt, welches aufgrund von lokalen Änderungen der Bewegungsrichtung (Bewegungscontrast) Szenen segmentieren kann.

Im vierten Kapitel dieser Arbeit wird das entwickelte neuronale Netzwerk in Bezug zu psychophysischen und neurophysiologischen Untersuchungen gebracht. Ferner wird das hier entwickelte Modell mit Modellen anderer Autoren verglichen.

2 Grundlagen

2.1 Psychophysische Grundlagen

Das visuelle System ist in der Lage, lokal unterschiedliche Bewegungen zu einem Perzept zusammenzubinden und eine Szenensegmentierung vorzunehmen. Dabei löst es scheinbar mühelos grundlegende Probleme der Bewegungsanalyse: z.B. das Apertur- und das Korrespondenzproblem. Anhand psychophysischer Untersuchungen lassen sich Vorhersagen über die zugrundeliegenden neuronalen Verschaltungen machen. Psychophysische Effekte liefern dabei Anhaltspunkte, wie das visuelle System zu einer eindeutigen, aber im physikalischen Sinne nicht unbedingt richtigen Bewegungsschätzung gelangt.

Im folgenden wird der Ausdruck „Bewegung“ synonym für kontinuierliche und scheinbare Bewegung („*apparent motion*“, „*Φ-movement*“) verwendet. Eine Bewegung wird scheinbar wahrgenommen, wenn Objekte sich nicht gleichmäßig von einem Ort zum anderen bewegen, sondern springen. Dies ist immer der Fall, wenn die Zeit diskretisiert wird, wie beispielsweise in Computersimulationen oder bei Kameraaufnahmen. Untersuchungen zu scheinbaren Bewegungen führten z.B. Julesz (1960 und 1971), Anstis (1970) sowie Braddick (1974) durch.

2.1.1 Apertur- und Korrespondenzproblem

Das Korrespondenzproblem resultiert aus der Unmöglichkeit, Objekte oder Teile eines Objektes in aufeinanderfolgenden Bildern in Übereinstimmung zu bringen. Die Abbildung 2.1 veranschaulicht das Korrespondenzproblem anhand zweier Flecken, die ihre Position vom Zeitschritt t_1 zu Zeitschritt t_2 ändern. Aufgrund fehlender Korrespondenz ist es nicht möglich festzustellen, welcher der Flecken wohin gesprungen ist.

Wird ein sich in eine beliebige Richtung bewegendes Balkenreiz durch eine runde Apertur beobachtet (Abbildung 2.2), so wird eine Bewegung orthogonal zu seiner Orientierung wahrgenommen. Die physikalische Bewegungsrichtung kann hiervon jedoch um bis zu ± 90 Grad abweichen. Dies ist das sogenannte Aperturproblem (Wallach 1935). In Kapitel 2.1.2.4 wird ferner gezeigt, daß die Form der Apertur die Bewegungswahrnehmung beeinflussen kann. Dies gilt auch für sich mit dem Balkenreiz mitbewegende markante Stellen, wie z.B. Ecken, innerhalb der Apertur.

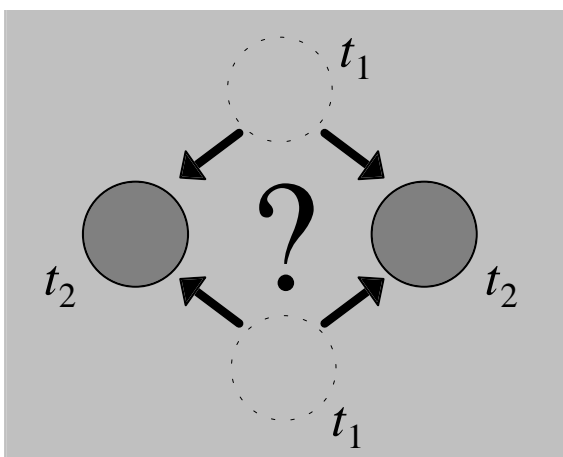


Abb. 2.1: Korrespondenzproblem. Zwei Flecken ändern ihre Position vom Zeitschritt t_1 (leere, gestrichelte Kreise) zu Zeitschritt t_2 (grau ausgefüllte Kreise). Welcher Fleck wechselte wohin? Aufgrund fehlender Korrespondenz bleibt diese Frage offen. Abbildung nach Gengerelli (1948).

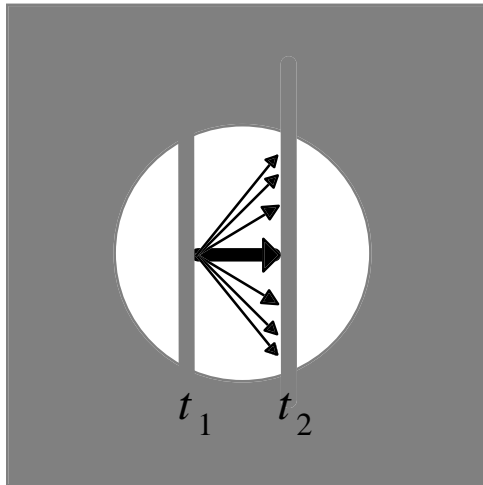


Abb. 2.2: Aperturproblem. Ein in beliebiger Richtung bewegter Balkenreiz scheint sich orthogonal zu seiner Orientierung zu bewegen (dicker Pfeil), wenn er durch eine runde Apertur beobachtet wird. Die physikalische Bewegungsrichtung kann nicht genau bestimmt werden. Sie kann um bis zu ± 90 Grad von der wahrgenommenen Bewegungsrichtung abweichen (dünne Pfeile). Abbildung nach Wallach (1935).

Das Aperturproblem steht im engen Zusammenhang mit dem Korrespondenzproblem. Die lokale Bewegung am Balkenreiz ist nicht eindeutig. Dennoch gelangt das visuelle System zu einer eindeutigen Bewegungsbestimmung. Es bleibt die Frage offen, welche neuronalen Verschaltungen hierfür verantwortlich sind. Die im nachfolgenden Kapitel 2.1.2 aufgeführten psychophysischen Untersuchungen geben Hinweise auf solche neuronalen Mechanismen.

2.1.2 Was erzeugt eine eindeutige Bewegungswahrnehmung?

2.1.2.1 Korrespondenz durch Objekteigenschaften

Das in Abbildung 2.1 dargestellte Korrespondenzproblem kann minimiert werden, wenn die beiden Flecken unterschiedliche Eigenschaften haben oder unterschiedlich weit springen. Nach Braddick (1974, Burt und Sperling 1981) wird Korrespondenz zwischen Flecken zweier aufeinanderfolgender Zeitschritte so hergestellt, daß die Sprungweiten minimal sind. Hierin ist das Gestaltprinzip der Nähe im übertragenen Sinne zu erkennen. Benachbarte Flecken oder Punkte gehören mit größerer Wahrscheinlichkeit zusammen als weiter entfernte.

Korrespondenz läßt sich aber auch durch Objekteigenschaften wie räumliche Frequenz, Polarität, Intensität, Farbe, Form und Orientierung erzeugen. Die räumliche Frequenz läßt sich sehr gut zur Herstellung von Korrespondenz verwenden (Green 1986). Dies deutet darauf hin, daß eine getrennte Verarbeitung von unterschiedlichen Sprungweiten in verschiedenen Frequenzbändern erfolgt (Bishop und di Lollo 1990). Einen starken Hinweis auf Korrespondenz liefert die Polarität des Intensitätskontrastes (Shechter und Hochstein 1990). Es gibt kein Korrespondenzproblem mehr, wenn die beiden Flecken der Abbildung 2.1 eine unterschiedliche Polarität zum Hintergrund aufweisen (Abb. 2.3). Die Intensität alleine ist nur bei genügend großem Unterschied zwischen den beiden Flecken hilfreich (Shechter und Hochstein 1990). Farbe (Simpson 1990, Ohtani et al. 1993) und Orientierung (Green 1986) sind nur schwache Merkmale für Korrespondenz in der Bewegungsverarbeitung.

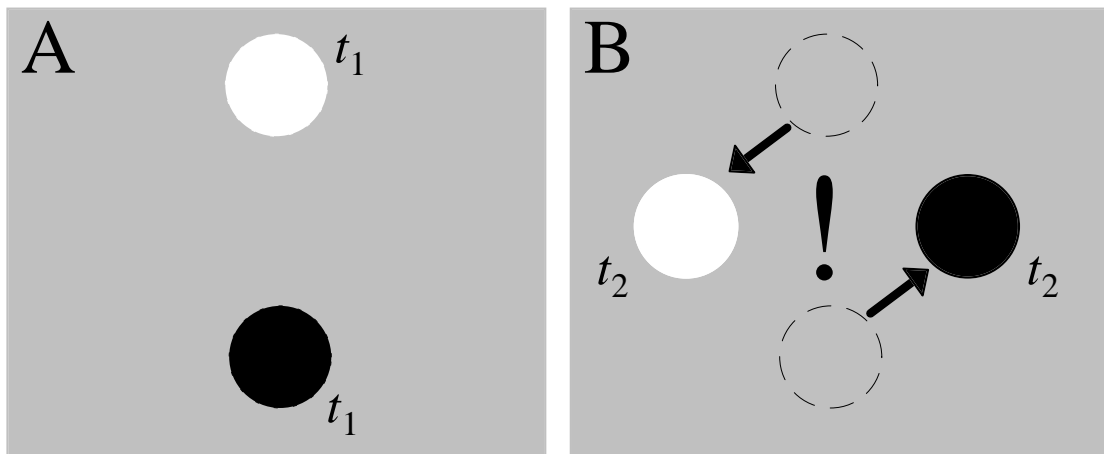


Abb. 2.3: Polarität stellt Korrespondenz her. A zeigt zwei Flecken mit unterschiedlicher Polarität zum Hintergrund zum Zeitpunkt t_1 . In B sind die Flecken zum Zeitpunkt t_2 dargestellt. Es ist kein Problem festzustellen, welcher Flecken wohin gesprungen ist. Die Polarität ist ein starker Hinweis, um eine Korrespondenz zwischen den Flecken beider Bilder herstellen zu können (nach Shechter und Hochstein 1990).

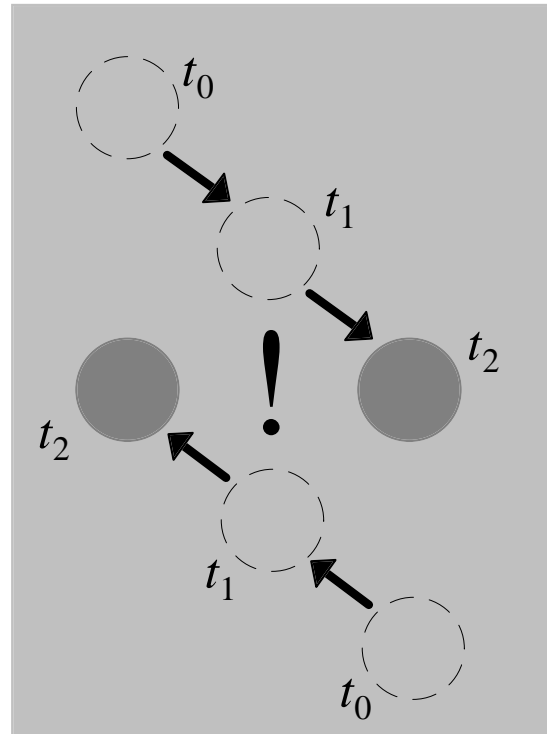
2.1.2.2 Korrespondenz durch zeitliche Integration der Bewegung

Das Korrespondenzproblem aus Abbildung 2.1 kann aber selbst dann behoben werden, wenn die Flecken identisch sind. Eine Möglichkeit besteht darin, die Vergangenheit des einzelnen Flecks, d.h. seine bisherige Bewegungsrichtung, mit in die Bewegungsverarbeitung einzubeziehen (Ramachandran und Anstis 1983, Anstis und Ramachandran 1987, Snowden und Braddick 1989). Hierbei wird davon ausgegangen, daß reale Objekte eine Trägheit besitzen und größtenteils keine schnellen Bewegungsänderungen auftreten. Dadurch wird es möglich, die weitere Bewegung vorherzusagen. Dies wird üblicherweise mit *priming* bezeichnet. Die Abbildung 2.4 zeigt hierfür ein Beispiel.

Dieses Experiment unterscheidet sich von den vorherigen um einen zusätzlichen Zeitschritt. Die Flecken werden auf die bekannte Konstellation, wie in Abbildung 2.4 gezeigt, hinbewegt, wodurch die Bewegungswahrnehmung eindeutig wird. Die eindeutige Bewegung aus Zeitschritt t_1 (Korrespondenz durch Nähe) wird verwendet, um eine eindeutige Bewegung in Zeitschritt t_2 zu erzielen. Wenn die vergangene Bewegung der Flecken keinen Einfluß auf die Bewegungsverarbeitung hätte, dann könnte das visuelle System keine Korrespondenz herstellen und es würde keine eindeutige Bewegungswahrnehmung vorliegen.

Die zeitliche Integration scheint auch für die Extraktion einer gleichförmig, geradlinigen Bewegung eines einzelnen Fleckes in sich zufällig bewegenden Fleckenmustern, den sogenannten *random dot patterns*, verantwortlich zu sein (Grzywacy et al. 1995). Hierbei ist die gegenüber dem Einzelneuron deutlich größere Integrationszeit (mehrere 100 ms) zu beachten (Watamaniuk et al. 1995).

Abb. 2.4: Korrespondenz durch zeitliche Integration I. Dargestellt sind die Positionen zweier Flecken zu drei aufeinanderfolgenden Zeitschritten t_0 , t_1 und t_2 . Die (scheinbare) Bewegung der beiden Flecken verläuft geradlinig. Die Konstellation zu den Zeitschritten t_1 und t_2 entspricht der der Abbildung 2.1. Die Wahrnehmung wird durch das Hinzufügen der Fleckenbewegung von Zeitschritt t_0 zu Zeitschritt t_1 eindeutig. Die bisherige Bewegung wird in die weitere Bewegungsverarbeitung miteinbezogen. Ohne eine zeitliche Integration wäre das Ergebnis der Bewegungsanalyse von Zeitschritt t_1 nach Zeitschritt t_2 mehrdeutig. Abbildung modifiziert nach Anstis und Ramachandran (1987).



Die Beeinflussung der Bewegungswahrnehmung von der bisherigen Bewegung wird auch in den Untersuchungen von Pinkus und Pantle (1997) deutlich. Ein Gitter, zum Beispiel mit sinusförmiger Helligkeitsverteilung, wird vom Zeitschritt t_1 zum Zeitschritt t_2 um 180° versetzt (Abb. 2.5 A). Die Bewegung ist somit mehrdeutig. Es ist nicht entscheidbar, ob sich das Gitter nach links oder rechts bewegt hat. Wird das Gitter aber zunächst um eine kleinere Phase verschoben und erst danach ein 180° -Phasensprung vollzogen, dann ist die Bewegungswahrnehmung eindeutig und entspricht der des ersten Schrittes (Abb. 2.5 B).

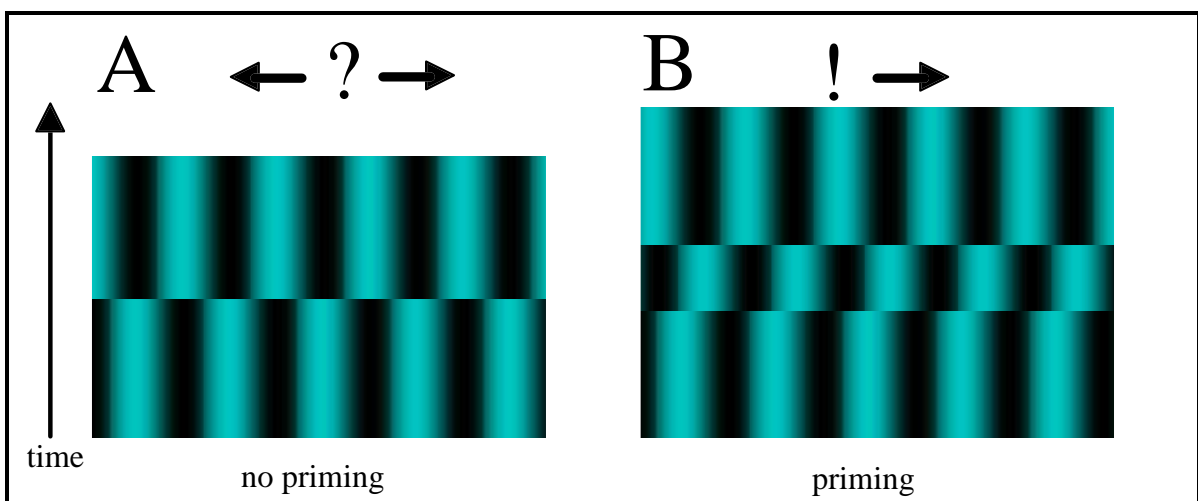


Abb 2.5: Korrespondenz durch zeitliche Integration II. **A:** Dargestellt ist ein Gitter, das um 180° in der Phase verschoben wird. Die Bewegungsrichtung ist mehrdeutig. **B:** Das gleiche Gitter vollzieht zunächst eine kleine Phasenverschiebung (hier 90°). Danach wird es wiederum um 180° in der Phase versetzt. Die Bewegungswahrnehmung ist eindeutig und wird vom vorherigen kleinen Schritt bestimmt (*priming*). Abbildung modifiziert nach Pinkus und Pantle (1997).

2.1.2.3 Korrespondenz durch räumliche Integration der Bewegung

Neben der zeitlichen Integration der Bewegungsinformation gibt es Hinweise auf eine räumliche Integration. Die räumliche Integration ist, wie viele psychophysische Untersuchungen zeigen, vielfältig und involviert sowohl verschiedene Auflösungen als auch besonders markante Merkmale wie Linienenden und Ecken. In diesem Zusammenhang werden die Termini „*motion capture*“ und „*motion assimilation*“ benutzt. Da es zahlreiche Untersuchungen zu Linienenden und Ecken, sogenannte Terminatoren, gibt, werden diese gesondert im nachfolgenden Kapitel 2.1.2.4 beschrieben.

Das in Abbildung 2.1 dargestellte Korrespondenzproblem läßt sich nicht nur durch zeitliche Integration, sondern auch durch eine räumliche Integration beheben (z.B. Yuille und Grzywacz 1988, Nawrot und Sekuler 1990 sowie Smith et al. 1994). In Abbildung 2.6 ist ein Beispiel für eine solche räumliche Integration nach Chang und Julesz (1984) sowie Yuille und Grzywacz (1988) dargestellt. Der untere Fleck der Konstellation aus Abbildung 2.1 erhält noch einen weiteren Flecken an seine Seite, der sich nach rechts oben bewegt. Diese zusätzliche Bewegung führt zu einer eindeutigen Bewegungswahrnehmung. Die Bewegung des unteren Fleckes scheint von der Bewegung des zusätzlichen Fleckes assimiliert zu werden.

Hierin spiegelt sich ebenfalls das Gestaltprinzip der Nähe wider. Benachbarte Flecken gehören wahrscheinlich zusammen, also haben sie wahrscheinlich auch die gleiche Bewegungsrichtung.

Räumliche Integration der Bewegungsinformation muß sich aber nicht auf eine Auflösung beschränken, wie zum Beispiel die Untersuchungen von Ramachandran und Cavanagh (1987) zeigen. Sie überlagerten sich in stochastische Richtungen bewegend Fleckenmuster mit einem sinusförmigen Gitter niedriger Auflösung und stellten fest, daß bei ausreichend großer Geschwindigkeit des Gitters die wahrgenommene Bewegungsrichtung der Flecken der des Gitters angeglichen wurde. Dies galt auch für den Fall einer kohärenten physikalischen Bewegung der Flecken in die dem Gitter entgegengesetzte Bewegungsrichtung. Die Bewegungsrichtungsdetektion niedriger räumlicher Frequenzen scheint nach diesen Untersuchungen die der hohen räumlichen Frequenzen zu dominieren.

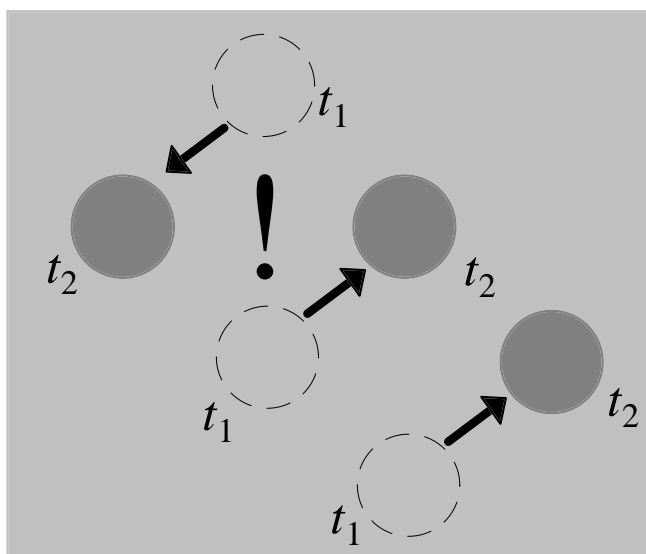


Abb 2.6: Korrespondenz durch räumliche Integration. Die Untersuchung unterscheidet sich von der in Abbildung 2.1 dargestellten durch einen weiteren, rechts unten dargestellten Fleck, der sich nach rechts oben bewegt. Durch diese zusätzliche Bewegung wird die Bewegungswahrnehmung eindeutig. Abbildung modifiziert nach Yuille und Grzywacz (1988).

In diese Art der räumlichen Integration paßt auch ein weiterer von MacKay (1961) sowie Yuille und Grzywacz (1988) untersuchter psychophysischer Effekt. Wird eine Schleife über ein sich zufällig bewegendes Fleckenmuster gezogen, so werden die Flecken als sich in Richtung der Schleifenbewegung bewegend wahrgenommen.

2.1.2.4 Einwirkung der Terminatoren auf die Bewegungswahrnehmung

Ein weiterer interessanter Effekt zeigt sich, wenn ein sich bewegendes Gitter durch eine rechteckige Apertur beobachtet wird (Abb. 2.7). Die wahrgenommene Bewegungsrichtung entspricht nicht der physikalischen (rechts). Die Gitter scheinen sich vielmehr parallel zu der längeren Seite der Apertur zu bewegen. Dieser Effekt wird als *barberpole illusion* bezeichnet. (Der Name geht auf rotierende Zylinder mit aufgedruckter Helix zurück, die vor amerikanischen Friseurgeschäften oft zu finden sind.) Hildreth (1984) sowie Nakayama und Silverman (1988a, b) machen die Linienenden, von ihnen „*Terminatoren*“ genannt, für diese Bewegungswahrnehmung verantwortlich. Die Bewegung des Balkenreizes ist aufgrund des Korrespondenzproblems mehrdeutig, die Bewegung der Linienenden hingegen nicht. Da an der langen Seite des Rechtecks sich mehr Linienenden als auf der kurzen Seite befinden, wird eine Bewegung in Richtung der kurzen Seite der Apertur wahrgenommen.

Shimojo et al. (1989) untersuchten die *barberpole illusion* stereoskopisch und fanden heraus, daß der Effekt verschwindet, wenn die Gitterbalken als hinter der Apertur befindlich wahrgenommen werden. Daraus folgerten sie die Existenz von zwei Arten von Linienenden: intrinsische und extrinsische Linienenden. Intrinsische Linienenden gehören zum Objekt, während extrinsische aus Verdeckungen entstehen. Im Gegensatz zu intrinsischen Linienenden bewirken extrinsische keinen „Barberpole-Effekt“, d.h. extrinsische Linienenden werden nicht zur Lösung des Aperturproblems herangezogen.

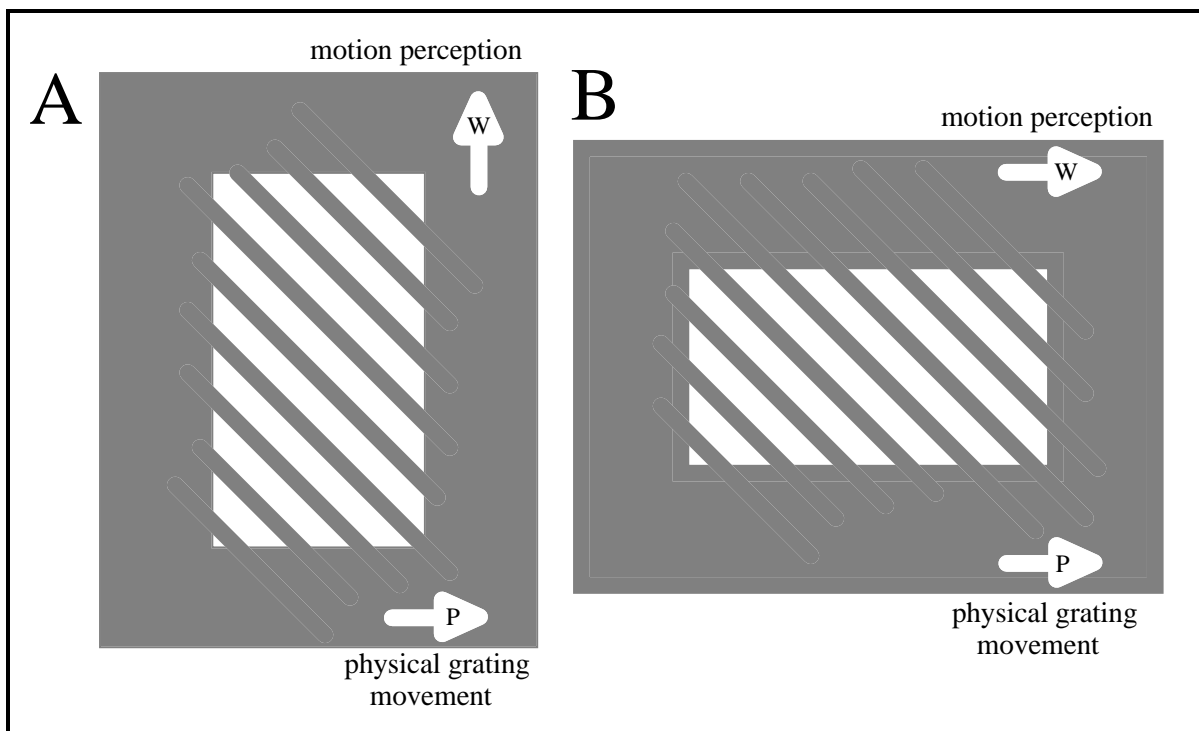


Abb. 2.7: „Barberpole Illusion“. Gitter mit gleicher Orientierung werden physikalisch nach rechts (Pfeile unten rechts mit P) bewegt und durch zwei unterschiedlich orientierte, rechteckige Aperturen (A, B) beobachtet. Die wahrgenommene Bewegungsrichtung (Pfeile oben rechts mit W) wird durch das Seitenverhältnis der Apertur bestimmt. Abbildung nach Wallach (1935).

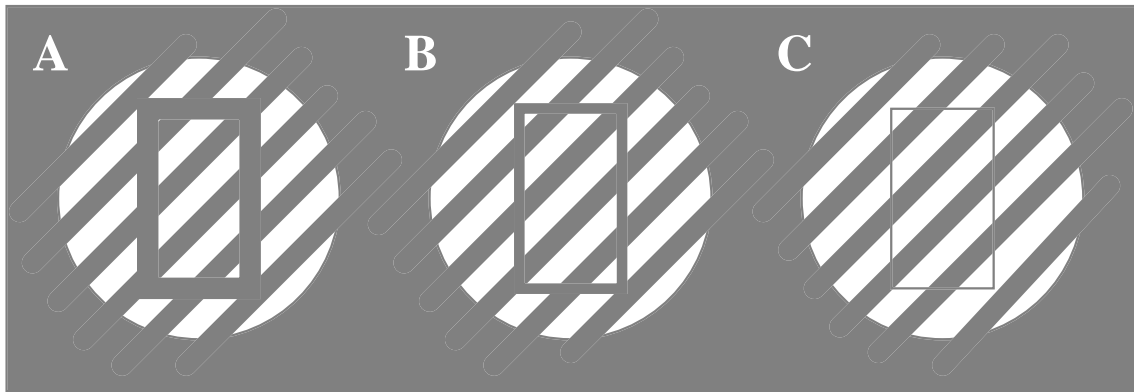


Abb. 2.8: Relative Größe als Tiefenhinweis zerstört „barberpole illusion“. Dargestellt sind drei verschiedene Versuchsanordnungen, jeweils bestehend aus einem sich bewegenden Gitter, einer runden Apertur und einem Rechteck innerhalb der Apertur, dessen Dicke im Verhältnis zur Dicke der Gitterbalken variiert wird. Die *barberpole illusion* wird bei **A** im Rechteck wahrgenommen, während sie bei **B** und **C** seltener bzw. nicht auftritt. Abbildung nach Vallortigara und Bressan (1991).

Unterstützt wird die Differenzierung der Linienenden auch durch die Untersuchungen von Vallortigara und Bressan (1991). Sie konnten ebenfalls zeigen, daß Tiefenhinweise Einfluß auf die *barberpole illusion* haben.

Wird zusätzlich zu einem durch eine runde Apertur beobachtetes, sich bewegendes Gitter ein stehendes Rechteck hinzugefügt (Abb. 2.8), dann tritt der „Barberpole-Effekt“ im Rechteck nur dann auf, wenn die Dicke des Rechtecks der Dicke der Gitterbalken entspricht. Rechteck und Gitter befinden sich in einer Tiefe, die Linienenden sind intrinsisch. Ist das Rechteck dünner als die Gitterbalken, dann wird es aufgrund seiner relativen Größe hinter dem Gitter gesehen, was extrinsische Linienenden erzeugt.

Kooi (1993) konnte durch die Verwendung von Aperturen mit glatten und eingeschnittenen Kanten zeigen, daß nicht nur das Seitenverhältnis der Apertur, sondern auch die lokale Orientierung der Aperturkanten die Bewegungswahrnehmung beeinflussen. Eine Apertur mit eingeschnittenen Kanten ist schematisch in Abb. 2.9 dargestellt. Die Einkerbungen führen zu keiner Änderung des Seitenverhältnisses der Apertur, aber die Linienenden bewegen sich nun nicht mehr vertikal und horizontal, sondern diagonal entlang der Aperturkanten. Sind die Einkerbungen groß genug ($\frac{1}{4}$ der Gitterperiode), tritt keine *barberpole illusion* auf, d.h., es wird eine diagonale Bewegung wahrgenommen. Diese Untersuchung unterstreicht die dominante Rolle der lokalen Bewegung der Linienenden bei der Bewegungswahrnehmung.

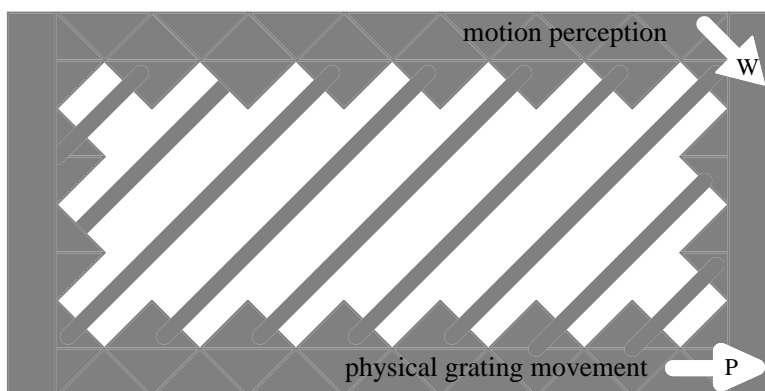


Abb. 2.9: Eingeschnittene Apertur stört „barberpole illusion“. Obwohl die Apertur horizontal ausgedehnt ist, wird bei einer physikalischen Gitterbewegung nach rechts eine Bewegung orthogonal zu den Linienorientierungen wahrgenommen. Abbildung nach Kooi (1993).

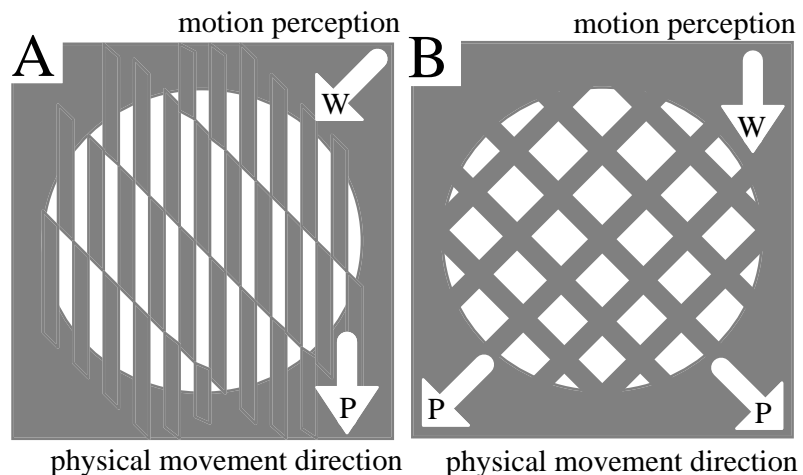
Die lokalen Bewegungen der Linienenden führen aber nicht immer zu einer entsprechenden globalen Bewegungswahrnehmung des Objektes. Wilson und Mast führten (1993) Untersuchungen mit bewegten Texturgrenzen durch und entdeckten dabei eine weitere interessante Bewegungssillusion. In Abb. 2.10 A ist als Beispiel ein Stimulus ihrer Untersuchung dargestellt. Er besteht aus um 180 Grad versetzten, vertikal angeordneten Gittern, deren Linienenden Scheinkonturen bilden. Werden diese Gitter physikalisch nach unten bewegt - die lokale Bewegung der Linienenden ist somit auch nach unten gerichtet - so wird eine Bewegung orthogonal zur Scheinkontur wahrgenommen. Dieser Effekt ist im fovealen Bereich stark, peripher hingegen schwach. Die Illusion ist außerdem abhängig von der Stimulusdauer. Kurze Präsentationen (≈ 60 ms) führen zu einer Bewegungswahrnehmung in Richtung der physikalischen Bewegungsrichtung, während große (> 150 ms) die Illusion hervorrufen. Offensichtlich kann bei einer längeren Darbietungsdauer aus den Linienenden eine Scheinkontur erzeugt werden, deren Bewegung, und nicht die der Linienenden, dann analysiert wird.

Daß die Einwirkung der Terminatoren auf die Bewegungsverarbeitung nicht nur davon abhängt, ob sie extrinsisch oder intrinsisch sind, zeigten Lorenceau und Shiffrar (1992). Sie wiesen eine deutliche Minderung des Einflusses der Terminatoren nach, wenn die Apertur zufällig „zerfranst“ wird oder die Exzentrizität ($> 7^\circ$) zunimmt. Im ersten Fall ist die Terminatorbewegung unzuverlässig und im zweiten Fall fehlt aufgrund der größeren rezeptiven Felder die Genauigkeit. Die Autoren folgerten daraus, daß nur sichere Terminatorbewegungen in die Bewegungsverarbeitung einfließen. Eine Verringerung des Intensitätskontrastes der Terminatoren zeigte ebenfalls einen Rückgang des Einflusses der Terminatorbewegung in der Bewegungsverarbeitung. Hieraus entnahmen sie, daß die Verarbeitung der Kantenbewegung und der Terminatorbewegung in zwei unterschiedlichen Verarbeitungsstufen erfolgen muß.

Die Abbildung 2.10 B zeigt eine weitere Illusion, die sogenannte *barberplaid illusion* (Adelson und Movshon 1982). Zwei Gitter werden orthogonal zu ihrer Orientierung in einer runden Apertur bewegt. In Abhängigkeit von ihren Unterschieden in räumlicher Frequenz, Orientierung, Geschwindigkeit, Farbe, Helligkeit und binokularer Disparität können die Gitter als zwei einzelne Objekte mit unterschiedlichen Bewegungsrichtungen oder als ein Objekt wahrgenommen werden (z.B. Stone et al. 1990, Trueswell und Hayhoe 1993, Heeley und Buchanan-Smith 1994, Alais et al. 1994, 1996 a, b). Die Bewegungsrichtung ist wie bei der *barberpole illusion* abhängig von der Ausdehnung der Apertur (Beutter et al. 1996). Für die Bewegungswahrnehmung könnten wiederum die durch die Überschneidungen entstandenen Terminatoren verantwortlich sein (Vallortigara und Bressan 1991).

Abb. 2.10 Scheinkonturbewegung und „barberplaid illusion“.

A: Die Linien werden physikalisch nach unten bewegt. Es entsteht aber ein Bewegungseindruck orthogonal zu den von den Linienenden hervorgerufenen Scheinkonturen (Peterhans et al. 1986). Abbildung nach Wilson und Mast (1993). **B:** Zwei Gitter werden orthogonal zu ihrer Orientierung bewegt. Wahrgenommen wird eine Bewegung nach unten.



2.1.3 Szenensegmentierung durch Bewegungsinformation

Die räumliche Integration der Bewegungsinformation ist beschränkt. Schließlich kann eine Szenensegmentierung aufgrund Bewegungsinformation sehr gut erfolgen. Dies gilt ebenso für räumlich abgegrenzte Objekte wie für sich überlappende, transparente Fleckenmuster. Die räumliche Integration muß folglich beschränkt sein. Nachfolgend sind psychophysische Experimente, die Hinweise auf neuronale Verschaltungen zur Szenensegmentierung aufgrund Bewegungsinformation geben, dargestellt. Einen Überblick über Segmentation und Integration in der visuellen Bewegungsverarbeitung liefert Braddick (1993).

Werden zwei zufällige Fleckenmuster überlagert, dann sind sie solange ununterscheidbar, bis sie in unterschiedliche Richtungen bewegt werden. Dieser Effekt reflektiert das Gestaltprinzip des gemeinsamen Schicksals (Ternus 1926, Koffka 1935). Die Flecken beider Muster erzeugen jeweils eine kohärente Bewegung, wodurch die beiden Muster voneinander segmentiert werden (z.B. Williams und Sekuler 1984, Williams et al. 1986). Die räumliche Integration scheint somit richtungs- und frequenzspezifisch zu sein. Ohne diese Einschränkungen wären *kinetic depth effect* und *structure from motion* (z.B. Husain et al. 1989, Treue et al. 1991, Treue et al. 1995 sowie Hildreth et al. 1995) nicht möglich.

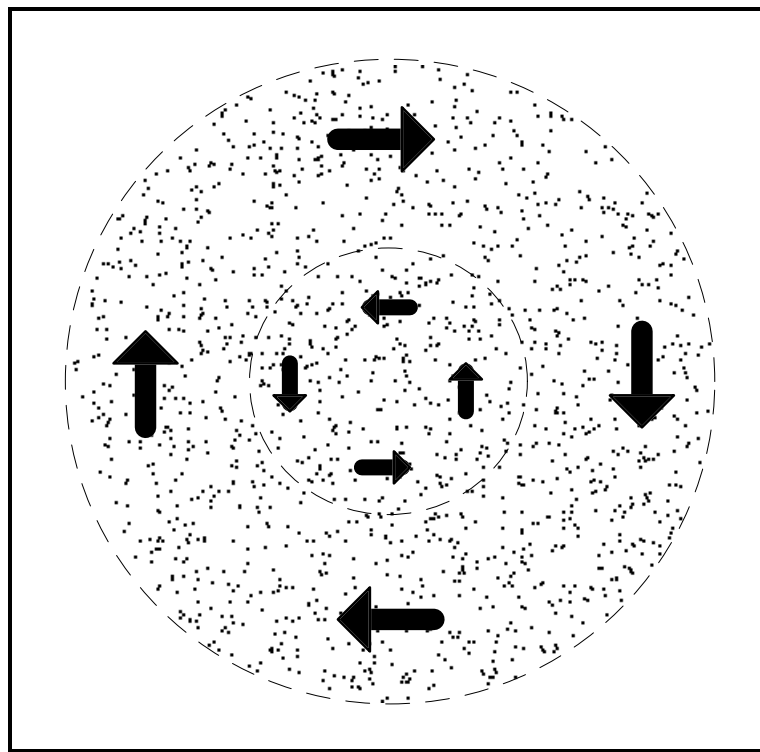


Abb. 2.11: Bewegungskontrast. Die Abbildung zeigt zwei Kreisscheiben. Die innere Scheibe dreht sich nach links, die äußere nach rechts. Die Scheiben bestehen nur aus Punkten. Die beiden gestrichelten Markierungen dienen nur der Visualisierung. Durch die gegenläufige Drehung ist deutlich eine Objektgrenze zwischen den Scheiben zu erkennen. An der Objektgrenze herrscht ein großer Unterschied in den Bewegungsrichtungen, ein hoher Bewegungs-kontrast.

Weitere psychophysische Untersuchungen zeigen, daß Objekte nicht aufgrund ihrer gemeinsamen Bewegungsrichtung zusammengebunden werden, sondern es vielmehr auf die lokalen Bewegungsrichtungsunterschiede, dem sogenannten Bewegungskонтраст, ankommt. In Abbildung 2.11 ist hierfür ein Beispiel nach Fenske (1995) und Fenske et al. (1995) dargestellt.

Der visuelle Reiz besteht aus zwei konzentrischen Kreisscheiben, die gegeneinander rotieren. Um die sichtbare Objektgrenze zwischen den beiden Scheiben wahrzunehmen, reicht eine Zusammenfassung aller Flecken mit gleicher Bewegungsrichtung nicht mehr aus. Werden aber die lokalen Bewegungsrichtungsunterschiede herangezogen, dann können die beiden Scheiben aufgrund des großen Bewegungsrichtungskонтраstes an der Grenze voneinander getrennt werden. Innerhalb der Kreisscheiben ist der Bewegungskонтраст nur gering.

Die Generierung von Objektgrenzen aufgrund lokaler Kонтраste gilt natürlich auch für andere Merkmale wie Intensität, Farbe und Textur. Weitere Beispiele für Formgenerierung durch Bewegung finden sich zum Beispiel in Nawrot et al. (1996) sowie Shipley und Kellman (1997).

2.2 Biologische Grundlagen

Unser visuelles System verschafft uns scheinbar mühelos einen dreidimensionalen Eindruck unserer Umgebung. Dazu nutzt es in Sekundenbruchteilen die Informationen aus, die es über die Retina aufnimmt. Neurophysiologische und anatomische Untersuchungen helfen uns, diese komplizierte Verarbeitung der visuellen Information - an dem viele Milliarden Neuronen beteiligt sind - besser zu verstehen. Zahlreiche Erkenntnisse stammen aus Untersuchungen an Katzen und Affen, insbesondere an Makaken, da deren Sehsystem in vielerlei Hinsicht dem des Menschen ähnelt.

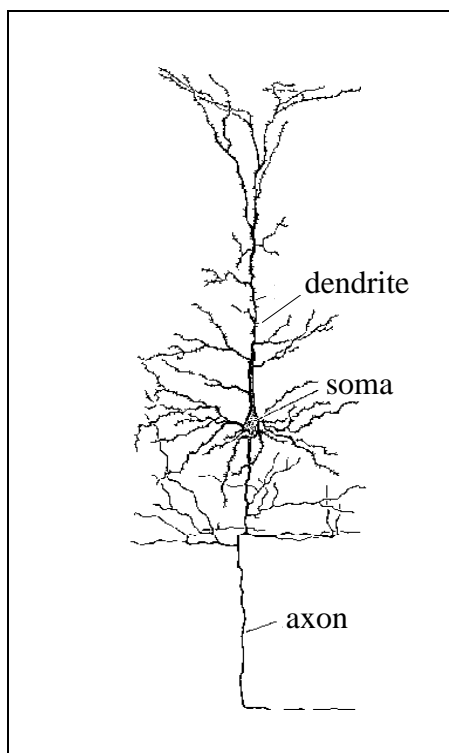
In diesem Grundlagenkapitel wird zunächst die Funktionsweise eines einzelnen Neurons beschrieben (Kapitel 2.2.1). Danach wird näher auf die Retina (Kapitel 2.2.2) und den für die visuelle Informationsverarbeitung wichtigen Teil des Kortex (Kapitel 2.2.3) eingegangen. Hierbei wird weniger Wert auf die Anatomie gelegt, sondern es steht vielmehr die Funktionalität der einzelnen Verarbeitungsbereiche im Vordergrund.

2.2.1 Das biologische Neuron

Das Neuron ist der elementare Baustein des Gehirns. Menschen besitzen im Zentralnervensystem etwa 100 Milliarden Neuronen, von denen jedes durchschnittlich mit 10 000 anderen über Synapsen verbunden ist. Ein einzelnes Neuron läßt sich in drei wesentliche Bestandteile untergliedern: Soma (Zellkörper), Dendritenbaum und Axon. Die Abbildung 2.12 zeigt ein solches typisches Neuron. Das Neuron empfängt Spannungsimpulse - die sogenannten Aktionspotentiale - in der Regel über seinen Dendritenbaum und verarbeitet diese. Bei ausreichend großer Aktivierung sendet es selbst über sein Axon Aktionspotentiale aus. Die Vorgänge der Aktionspotentialgenerierung und der synaptischen Signalübertragung werden im folgenden genauer beschrieben. Zuvor wird jedoch auf den grundlegenden Ionenaustausch an der Zellmembran eingegangen.

Die hauptsächlich aus Lipiden und Proteinen bestehende Membran grenzt Intra- und Extrazellularraum voneinander ab. Sowohl das Zellplasma als auch die extrazelluläre Flüssigkeit sind gute elektrische Leiter, da die in ihnen vorhandenen Salze weitgehend in Anionen und Kationen dissoziiert sind. Die für die Erregungsbildung wichtigsten Ionen sind K^+ , Na^+ , Ca^{2+} und Cl^- . Die Lipidschicht ist für diese Ionen impermeabel. Sie können die Membran nur über als selektive Kanäle fungierende Proteine durchqueren. Die Innenkonzentration der K^+ -Ionen ist um ein Vielfaches größer als deren Außenkonzentration. Im Gegensatz dazu verhalten sich die Innen- und Außenkonzentrationen von Na^+ , Ca^{2+} - und Cl^- umgekehrt, woraus eine elektrische Potentialdifferenz, das Membranpotential, resultiert. Die unterschiedlichen Konzentrationen ergeben sich aus dem Zusammenwirken von passiven und aktiven Ionen-transporten. Der aktive Ionen-transport wird dabei von Ionenpumpen, insbesondere der Na^+ / K^+ -Pumpe, gewährleistet. Wegen des resultierenden Konzentrationsgradienten diffundieren die Ionen auch passiv über die Kanalproteine durch die Zellmembran. Dadurch entsteht ein elektrisches Feld zwischen Innen- und Außenraum, was den Diffusionsstrom schließlich restringiert. Hierbei ist das Potential nach Hodgkin und Katz (1949) von den jeweiligen Ionen-Permeabilitäten abhängig. Das Ruhemembranpotential liegt je nach Zelltyp zwischen -50 und -100 mV.

Nerven- oder Muskelzellen sind erregbare Zellen. Werden sie so stark gereizt, daß ein bestimmtes Potential, das Schwellenpotential, überschritten wird, dann ändert die Membran ihre Ionenleitfähigkeit. Dies gilt insbesondere für die Natrium- und Kaliumleitfähigkeit. Durch das Öffnen von weiteren Kanälen am Axonhügel fließen weitere Na^+ -Ionen in die Zelle hinein, wodurch ein positiver Rückkopplungsmechanismus - der *Hodgkin-Huxley-Zyklus* (Hodgkin und Huxley 1952) - entsteht, der einen starken Anstieg des Membranpotentials zur Folge hat (*Depolarisation*). Verzögert öffnen sich auch die K^+ -Kanäle, wodurch entsprechend des Diffusionsgradienten K^+ -Ionen in den Extrazellularraum fließen. Dies bewirkt zunächst ein Absinken des Membranpotentials unter seinen Ruhewert (*Hyperpolarisation*) und danach auf diesen zurück. Das Neuron kann unmittelbar danach kein weiteres Aktionspotential generieren. Dieses Verhalten wird mit absoluter Refraktärzeit bezeichnet. Daran schließt sich die relative Refraktärzeit, das ist der Zeitraum, in der das Neuron nur schwer erregbar ist, an.



Nervenzellen haben über Synapsen Kontakt mit anderen Nervenzellen. Es gibt zwei Arten von Synapsen: chemische Synapsen und elektrische Synapsen. Im folgenden wird die Signalübertragung der häufiger vorkommenden chemischen Synapse beschrieben.

Abb. 2.12: Biologisches Neuron. Abbildung nach Kuffler et al. (1984). Informationsfluß von oben nach unten.

Die präsynaptische Membran ist von der postsynaptischen durch einen Spalt getrennt. Das Ende des präsynaptischen Axons ist mit synaptischen Vesikeln gefüllt. Diese sind in das Aktinskelett der Nervenendigung eingebunden und beinhalten eine Transmittersubstanz. Das Eintreffen eines Aktionspotentials in den präsynaptischen Endknopf führt zu einer Depolarisation, die das Öffnen von spannungsabhängigen Ca^{2+} -Kanälen verursacht. Die einströmenden Ca^{2+} -Ionen und die Depolarisation sind die Voraussetzung für die dann stattfindende regulierte Exozytose. Dazu gehört der Transport der Vesikel an die für die Membranfusion spezialisierten Stellen der Plasmamembran, an denen die Vesikel andocken, und das Ausschütten des Transmitters in den synaptischen Spalt. Die Transmitter diffundieren danach zur postsynaptischen Membran, wo sie sich an spezifische, chemisch gesteuerte Rezeptormoleküle anlagern und eine Permeabilitätsänderung hervorrufen. Je nach Art der Transmittersubstanz und der Rezeptormoleküle kann das Aktionspotential depolarisierend oder hyperpolarisierend auf den postsynaptischen Dendriten wirken. Erregende Transmitter wie Acetylcholin und Glutamat erhöhen die Ionenleitfähigkeit der Na^+ -, K^+ - und Cl^- -Ionenkanäle. Durch den überwiegenden Einstrom an Na^+ -Ionen kommt es dann zu einer Depolarisation, dem exzitatorischen postsynaptischen Potential (EPSP). Im Gegensatz zu den erregenden Transmittern erhöhen hemmende Überträgerstoffe wie Glyzin und GABA (γ -Aminobuttersäure) nicht die Na^+ -Ionenleitfähigkeit, was zu einer Hyperpolarisation der Zelle führt. Der zeitliche Verlauf des postsynaptischen Potentials ist im Vergleich zum Aktionspotential graduiert und zeitlich ausgedehnt.

2.2.2 Die Retina

Die Retina führt die erste Stufe der visuellen Informationsverarbeitung aus (einen Überblick liefern z.B. Kolb (1994) und Sterling et al. (1986)). Sie lässt sich anatomisch in acht gut unterscheidbare Schichten einteilen. Eine äußere und eine innere Membran begrenzen die Retina. Die Neuronen der Retina lassen sich in Rezeptoren, Bipolar-, Horizontal-, Amakrin- und Ganglienzellen einteilen. Die Photorezeptoren der Vertebratenretina, das sind die Stäbchen und Zapfen, bestehen aus einem photosensitiven Segment, das außerhalb der äußeren Grenzmembran liegt, und einem sich auf der Innenseite befindlichem Segment mit Zellkern. Diese Innenglieder bilden die äußere Körnerschicht. Die innere Körnerschicht setzt sich aus den Somata der Horizontal-, Bipolar-, und Amakrinzellen zusammen. Die beiden Körnerschichten werden durch die äußere plexiforme Schicht, in denen die Dendriten der Horizontal- und Bipolarzellen verlaufen, und die Henlesche Faserschicht getrennt. Die Schicht der Ganglienzellen wird durch die innere plexiforme Schicht von der inneren Körnerschicht separiert. Die Axone der Ganglienzellen bilden den Sehnerv. Morphologisch werden die Ganglienzellen in α -, β - und γ -Zellen eingeteilt (Boycott und Wässle 1974). Diese Einteilung beruht auf der unterschiedlichen Größe des Dendritenbaumes. α -Zellen haben im Gegensatz zu β -Zellen große Zellkörper und einen weit verzweigten Dendritenbaum. α - und β -Zellen werden in zwei Unterklassen eingeteilt. Zellen, die in Sublamina a der inneren plexiformen Schicht verzweigen, werden mit a-Zellen bezeichnet. Entsprechend werden Zellen, deren Dendriten in Sublamina b enden, b-Zellen genannt (Famiglietti und Kolb 1976).

Während Rezeptoren, Horizontal- und Bipolarzellen mit graduierten Potentialen auf Reizung reagieren, leiten Ganglienzellen Informationen über Aktionspotentiale weiter. Der Potentialverlauf der Amakrinzellen ist nicht eindeutig einzuteilen.

Funktionell weist die Retina verschiedene Subsysteme auf. Eine auf den beiden Rezeptortypen basierende Unterscheidung ist die Einteilung in ein photopisches und ein skotopisches Sehen. Für das photopische Sehen, das Sehen von Einzelheiten bei Helligkeit, werden die Zapfen, für das skotopische Sehen, das im Vergleich zum photopischen Sehen ungenauere Dämmerungssehen, die Stäbchen verwendet. Die räumliche Verteilung der Rezeptoren in der Retina sowie deren unterschiedliche Konvergenz auf nachfolgende Zellen führt zu dieser Funktionalität. So sind die etwa 6 Millionen Zapfen der menschlichen Retina zahlreich in der „Fovea centralis“ - dem „Ort des schärfsten Sehens“ -, aber nur sehr gering in der Peripherie vertreten. Die rund 120 Millionen Stäbchen befinden sich hingegen am häufigsten außerhalb der Fovea centralis (Übersicht z.B. in Wässle und Boycott (1991)).

Bipolarzellen erhalten ihre Eingangssignale entweder von Zapfen oder Stäbchen. Die Art der Verschaltung zwischen Rezeptor und Bipolarzelle führt zu zwei funktionellen Gruppen: den Bipolarzellen des ON-Pfades und denen des OFF-Pfades. Die ON-Bipolarzellen reagieren mit einer Depolarisation, wohingegen OFF-Bipolarzellen mit einer Hyperpolarisation auf eine Beleuchtung des Zentrums ihres rezeptiven Feldes antworten. Mit *rezeptivem Feld (RF)* einer Zelle - genau genommen das klassische rezeptive Feld (cRF) - wird der Bereich des Sehraumes bezeichnet, in dem visuelle Stimuli die Aktivierung der Zelle direkt beeinflussen. Das rezeptive Feld der Bipolarzellen besitzt eine antagonistische Zentrum/Umfeld-Struktur. Diese ist in der Verschaltung der Rezeptoren und Horizontalzellen zu Bipolarzellen begründet. Die Bipolarzellen sind exzitatorisch mit den retinalen Ganglienzellen verschaltet. ON- und OFF-Pfad bleiben hierbei anatomisch getrennt, da die Axone der ON-Bipolarzellen ausschließlich in Sublamina b und die der OFF-Bipolarzellen in Sublamina a münden. Die Bipolarzellen geben ihre Signale zusätzlich an die Amakrinzellen weiter. Über die Funktion der Amakrinzellen ist wenig bekannt, aber aufgrund ihrer Verschaltung zu Ganglienzellen, die der der Horizontal- zu den Bipolarzellen ähnelt, wird auch hier eine Zentrum/Umfeld-Struktur angenommen. Das ON- und OFF-Antwortverhalten retinaler Ganglienzellen der Katze auf Stimulation mit kleinen Lichtflecken, dargestellt in Abbildung 2.12, ist bereits seit Kuffler (1953) bekannt. Die Zentrum/Umfeld-Struktur bei retinalen Ganglienzellen des Affen geht auf Hubel und Wiesel (1960) zurück.

Eine weitere Unterteilung der retinalen Ganglienzellen geht auf ihre raum-zeitlichen Übertragungseigenschaften zurück. In der Katzenretina werden X-, Y-, und W-Zellen unterschieden (Enroth-Cugell und Robson 1966). Die β -Zellen sind hierbei funktionell dem X-System und die α -Zellen dem Y-System zuzuordnen (Cleland et al. 1975). Das zeitliche Antwortverhalten der X- als auch der Y-Zellen wird durch einen transienten (phasischen) und einen anhaltenden (tonischen) Anteil charakterisiert. Y-Zellen haben im Vergleich zu den X-Zellen einen deutlich stärkeren transienten Anteil, dafür aber keinen oder einen deutlich schwächeren anhaltenden Anteil. Y-Zellen scheinen deswegen besonders für die Verarbeitung von zeitlichen Transienten, wie sie zum Beispiel bei Bewegungen auftreten, und X-Zellen für die Verarbeitung stationärer Stimuli geeignet zu sein. Weiterhin unterscheiden sich die X-Zellen gegenüber den Y-Zellen durch kleinere rezeptive Felder und langsamerer Signalübertragung. Die RF-Größe steigt im allgemeinen mit zunehmender Exzentrizität (Katze: Boycott und Wässle (1974), Kolb et al. (1981), Übersicht in Wässle und Boycott (1991); Affe: Übersicht z.B. in Watanabe und Rodieck (1989)). Über die W-Zellen ist bisher wenig bekannt. Das funktionelle X-System der Katze entspricht dem *parvozellularen* und das Y-System dem *magnozellularen* Verarbeitungspfad des Affen. Die geringen Unterschiede zwischen der Makaken-Retina und der menschlichen Retina werden in Dacey und Petersen (1992) beschrieben. Im Gesamtmodell der Abbildung 1.1 spiegeln sich die Eigenschaften des X-Systems im Kontur/Form-System und die des Y-Systems im Transientensystem wider.

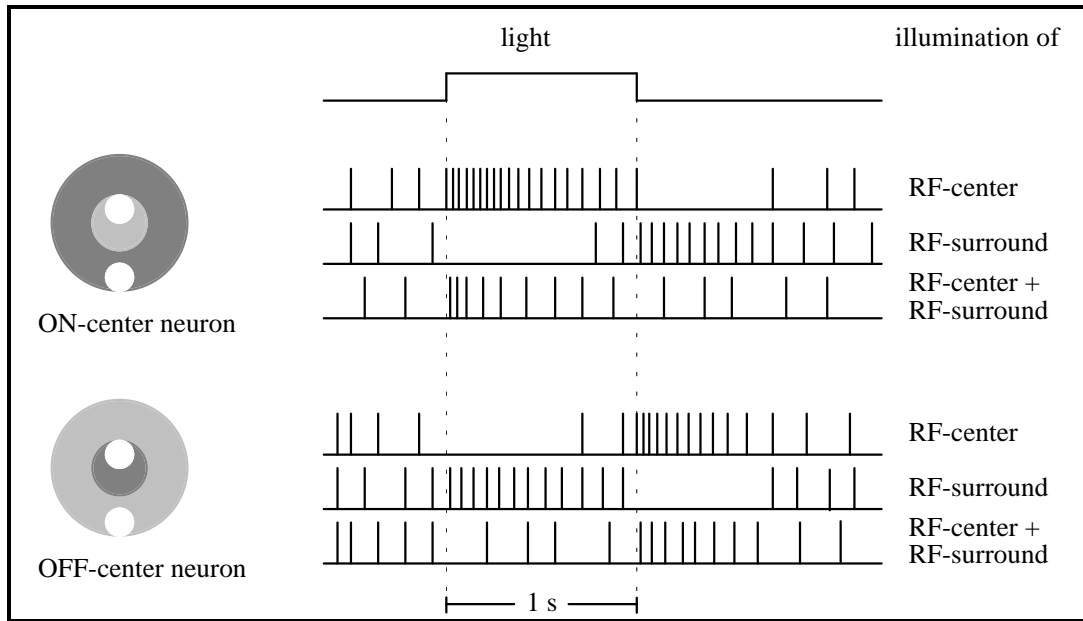


Abb. 2.13: Eigenschaften der ON-Zentrum- und OFF-Zentrum-Ganglienzellen. Links sind die RF einer ON- und einer OFF-Ganglienzelle dargestellt. Das ON-Gebiet ist hellgrau, das OFF-Gebiet dunkelgrau und der Lichtfleck weiß. Rechts ist die Aktivität der Ganglienzellen in Abhängigkeit von der Zeit für verschiedene Beleuchtungen schematisch aufgetragen (modifiziert nach Schmidt 1985).

2.2.3 Die visuellen Verarbeitungsbereiche des Kortex

Vom visuellen System des Makaken sind bisher 32 visuelle Areale sowie einige subkortikale Zentren, zu denen zum Beispiel die beiden *Retinae*, das *Corpus geniculatum laterale* (CGL oder englisch LGN) und der *Colliculus superior* gehören, bekannt. Die bereits in der Retina getrennten Verarbeitungspfade bleiben auch im CGL getrennt. Diese Trennung wird größtenteils in den visuellen Arealen des Kortex aufrechterhalten. Die Abbildungen 2.14 und 2.15 zeigen eine Übersicht der wichtigsten Areale und ihre mögliche Einordnung in die Verarbeitungspfade sowie den hierarchischen Aufbau der visuellen Informationsverarbeitung im Gehirn. Die Hauptaufgabe des Magno-Verarbeitungspfad (M-Pfad) ist die Bestimmung des Ortes und der Bewegung der Objekte. Im Parvo-Verarbeitungspfad (P-Pfad) ist hingegen die Klassifizierung der Objekte das primäre Ziel. Im folgenden werden die einzelnen Verarbeitungstufen näher vorgestellt, wobei die Bewegungsverarbeitung im Vordergrund stehen soll.

Die Axone der retinalen Ganglienzellen bilden den Sehnerv, der im CGL endet. Das CGL besteht beim Affen aus sechs Zellschichten. Die ersten beiden Schichten gehören zum M-Pfad, die anderen zum P-Pfad. Die Neuronen des CGL haben wie die retinalen Ganglienzellen konzentrisch angeordnete, antagonistische rezeptive Felder, wobei räumlich benachbarte CGL-Zellen auch räumlich benachbarte rezeptive Felder haben (Wiesel und Hubel 1966). Neuronen der unteren Verarbeitungstufen, die retinalen Ganglienzellen und die Zellen des CGL, antworten auf sich bewegende Objekte, haben aber keine Richtungsselektivität. Diese tritt erst bei Neuronen des ersten visuellen Areals auf (Hubel und Wiesel 1968).

Die visuelle Information wird im Kortex hierarchisch weiterverarbeitet. Das erste visuelle Areal wird beim Affen mit V1 bei der Katze mit Areal 17 bezeichnet. Auch im Kortex gibt es Neuronen, deren Antwortverhalten als primär transient oder anhaltend charakterisiert werden kann (Ikeda und Wright 1974). Das Areal V1 ist in sechs Schichten unterteilt. Die aus dem CGL stammenden Fasern enden meistens in den Schichten 4 und 6. Die kortikalen Neuronen haben andere Eigenschaften als die des CGL, so verfügen viele von ihnen über Orientierungs-

und Richtungsselektivität. Neuronen mit konzentrischem rezeptiven Feld kommen hauptsächlich in Schicht 4C vor. In den Schichten 4B, 4Ca und 6 sind etwa ein Drittel der Neuronen richtungsselektiv (Hawken et al. 1988). Sie werden dem M-Pfad zugeordnet, wohingegen die Schichten 4A und 4Cb dem P-Pfad angehören. Auf die Orientierungs- und Richtungsselektivität der Neuronen wird im folgenden näher eingegangen.

Bereits 1960 entdeckten Hubel und Wiesel bei einem Teil der kortikalen Neuronen eine Orientierungsselektivität. Diese Neuronen zeichnen sich durch ein länglich ausgedehntes rezeptives Feld aus, wodurch sie auf Stimulation mit Balkenreizen dieser Orientierung besonders stark aktiv werden. Solche rezeptiven Felder können durch die Verschaltung von CGL-Zellen mit kollinear angeordneten rezeptiven Feldern entstehen. Dabei kann die Verschaltung aus ON- und OFF-Zellen (Hubel und Wiesel 1965) oder ausschließlich aus ON- oder OFF-Zellen bestehen (Heggelund 1981). Neuronen, deren rezeptives Feld in ON- und OFF-Bereiche unterteilt sind, werden - auf Hubel und Wiesel zurückgehend - als *Simple-Zellen* bezeichnet. Fehlt eine solche Unterteilung, dann werden orientierungsselektive Neuronen mit *Komplex-Zelle* bezeichnet. Die Komplex-Zellen haben größere rezeptive Felder und antworten tolerant gegenüber kleinen Positionsverschiebungen. Ihre rezeptiven Felder können aus Simple-Zellen aufgebaut werden. Weiterhin sind Komplex-Zellen bekannt, deren rezeptive Felder noch einen zusätzlichen inhibierenden Endbereich beinhalten. Diese sogenannten *Hyperkomplex-Zellen* (oder englisch *End-Stopped-Cells*) reagieren bevorzugt auf Kantenenden.

Die Aktivität der Neuronen wird aber nicht nur durch Stimuli innerhalb ihres klassischen rezeptiven Feldes bestimmt. So gibt es zusätzliche laterale Kopplungen zwischen Neuronen, die ähnliche Merkmale kodieren. Neurophysiologische Untersuchungen von Eckhorn et al. (1988) zeigen eine starke Korrelation der Aktivitäten lateral gekoppelter Neuronen. Weitere Untersuchungen von Hirsch und Gilbert (1991) deuten auf modulatorische Interaktionen zwischen diesen Neuronen hin. Die laterale Kopplung scheint dabei von der kortikalen Entfernung der Neuronen und von ihrem Unterschied in der Orientierungspräferenz abhängig zu sein (Frien et al. 1996). Die Orientierungscharakteristik der Neuronen mit synchronisierten neuronalen Aktivitäten ist hierbei schärfer, als es aufgrund ihrer RF-Architektur zu erwarten wäre. In diesem Zusammenhang wird der Begriff des klassischen rezeptiven Feldes auf ein *Assoziations-Feld* (AF, Eckhorn et al. 1990) ausgedehnt.

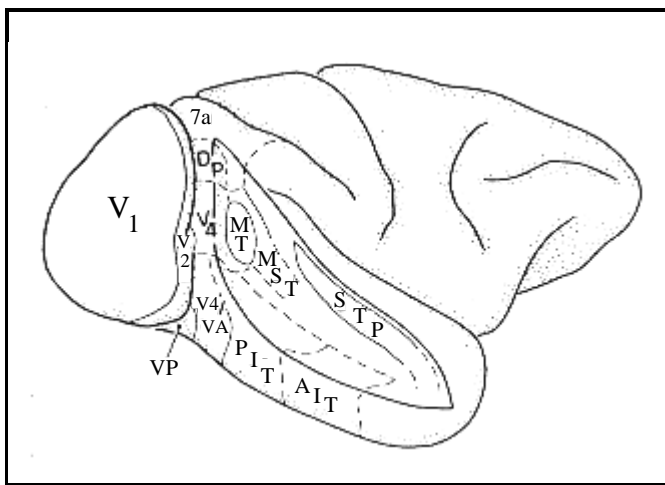


Abb. 2.14: Die Areale der visuellen Informationsverarbeitung. Dargestellt ist eine seitliche Ansicht des Makakenkortex mit geöffnetem superior temporalen sulcus, so daß die Areale MT und MST sichtbar werden. Die Augen wären in dieser Abbildung rechts. Abbildung aus Maunsell und Newsome (1987).

Der parvozelluläre Verarbeitungspfad wird über die weiteren visuellen Areale V2 und V4 zu dem inferotemporalen Kortex (IT) fortgesetzt. IT scheint eine wesentliche Rolle bei der Erkennung von Objekten zu spielen (z.B. Maunsell und Newsome 1987). IT-Neuronen, die ihre Eingabe hauptsächlich von V2 und V4 erhalten, scheinen die Merkmale Form, Farbe und Textur eines Objektes zu einem Perzept zu integrieren. Es gibt erste Hinweise darauf, daß die Antworten der IT-Neuronen mit der Wahrnehmung der Affen korrespondieren könnten (z.B. Nakamura et al. 1994).

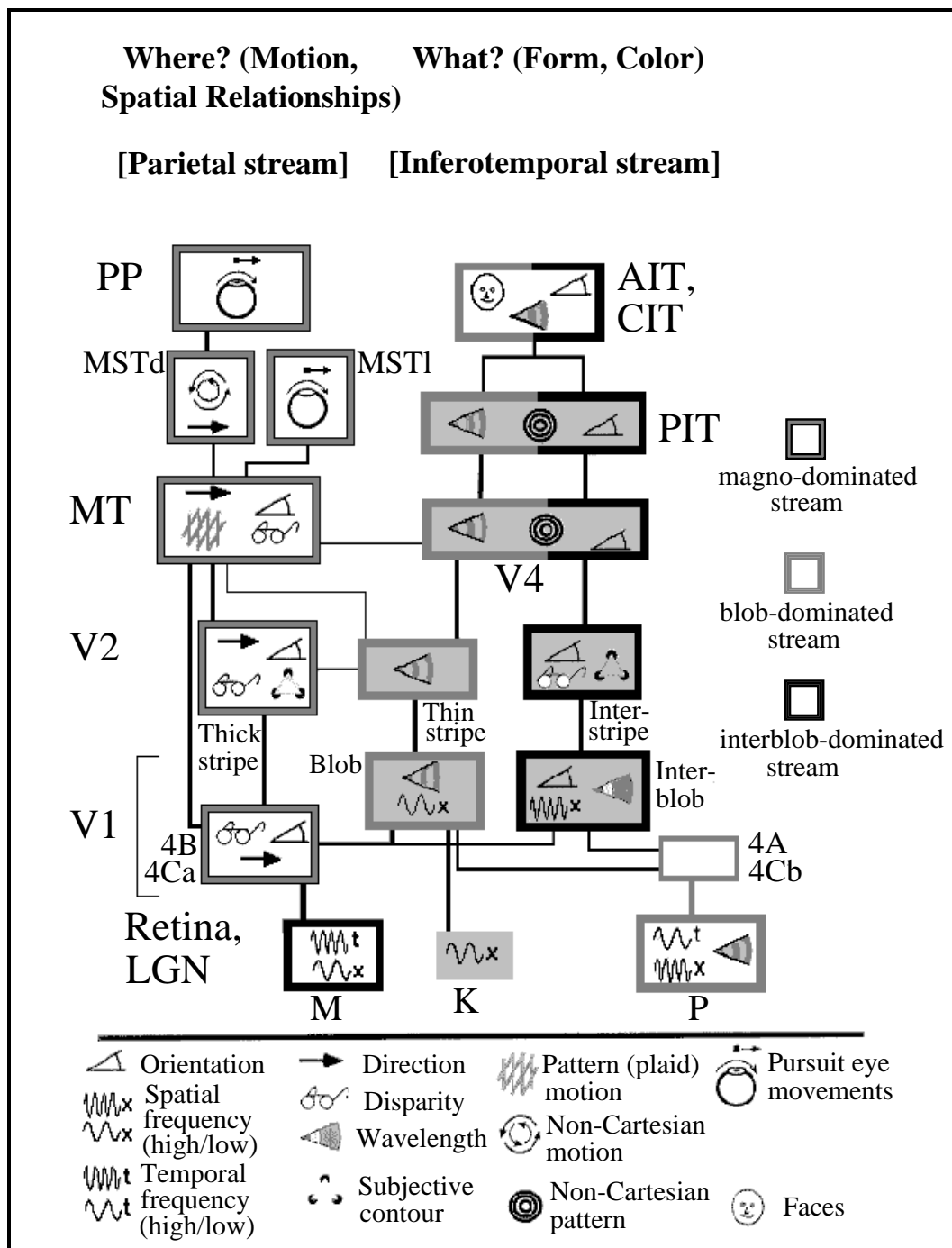


Abb. 2.15: Hierarchische Organisation der Verarbeitungspfade im Makaken. Die Abbildung zeigt schematisch die visuellen Areale und die Informationsverarbeitung im M- und P-Pfad. Abbildung nach van Essen und Gallant (1994).

Der magnozelluläre Verarbeitungspfad führt zu dem mediotemporalen Areal (MT nach Allman und Kaas (1971) oder V5 nach Zeki (1974)). MT ist spezialisiert für die Analyse der Stimulus Bewegung. Die MT-Neuronen erhalten ihre Eingabe primär aus den Schichten 4 und 4Ca (Maunsell und van Essen 1983, Fries et al. 1985). Ein sehr großer Teil der Neuronen in MT sind richtungs- und geschwindigkeitsspezifisch, d.h., ihre Aktivität ist für eine bestimmte Bewegungsrichtung bzw. Geschwindigkeit maximal (Dubner und Zeki 1971, Zeki 1974, Zeki 1978, Maunsell und van Essen 1983, Albright 1984). Diese Neuronen können durch sich nicht bewegendes Intensitätsmodulationen aktiviert werden; ihre Aktivität ist aber dann deutlich geringer als bei bewegtem Stimulus. Das in Abbildung 2.16 dargestellte Polardiagramm zeigt die gemittelte Aktivität eines MT-Neurons in Abhängigkeit von der Bewegungsrichtung des rechteckförmigen Stimulus (Maunsell und van Essen 1983).

MT beinhaltet ähnlich wie V1 eine kolumnare Organisation (Albright et al. 1984), d.h. die Richtungspräferenz benachbarter Neuronen innerhalb einer vertikalen Kolumne ändert sich nur gering und die der benachbarten Neuronen in angrenzenden Kolumnen ist entgegengesetzt. Dies gilt nach Wang et al. (1995) ebenso für das bei der Katze verantwortliche Areal zur Bewegungsdetektion PMLS (englisch: posteromedial lateral suprasylvian area).

Albright et al. (1984) zeigten, daß die Antworten von 55% der MT-Neuronen nur schwach abhängig von der Farbe, der Form, der Länge oder der Orientierung des Stimulus waren. 45% der Neuronen zeigten eine Abhängigkeit der Antwortstärke von der Ausdehnung des Stimulus bei gleichbleibender Richtungsselektivität. 84% der untersuchten Neuronen hatten eine eindeutige Richtungsselektivität und nur 8% ein von der Bewegungsrichtung unabhängiges Antwortverhalten.

Die Richtungsspezifität der MT-Neuronen beruht wahrscheinlich auf einem Inhibitions- und einem Fazilitationsmechanismus (Mikami et al. 1986). Die bereits in dem Kapitel 2.1 dargestellte Voraktivierung wurde auch durch neurophysiologische Untersuchungen an MT-Neuronen bestätigt. Es scheint hierbei sowohl fazilitatorische als auch inhibitorische Mechanismen zur Bewegungsdetektion zu geben (Movshon et al. 1978, Mikami et al. 1986). Mikami zeigte (1992) unter der Verwendung von Lichtblitzen als Stimuli am wachen Affen, daß es häufig ein starkes inhibitorisches Gebiet außerhalb des exzitatorischen rezeptiven Feldes einzelner MT-Neuronen gibt. Dazu nahm Mikami die Aktivität von MT-Neuronen bei Stimulation mit einem Lichtblitz innerhalb des exzitatorischen rezeptiven Feldes auf und verglich diese mit der Aktivität, die der Lichtblitz evoziert, wenn diesem ein weiterer Lichtblitz innerhalb einer bestimmten Zeitspanne (36 ms - 161 ms) an einem benachbarten Ort ($0,375^\circ$ - $2,1^\circ$ Sehwinkel) voranging. Mikami stellte fest, daß ein großer Teil der MT-Neuronen deutlich schwächer aktiv war, wenn die beiden Lichtblitze sich scheinbar in die der Richtungspräferenz des Neurons entgegengesetzten Richtung bewegten. Im Gegensatz dazu war bei einer scheinbaren Bewegung in Richtung der Richtungspräferenz keine Fazilitation signifikant zu erkennen. Dies änderte sich allerdings durch Hinzufügen von weiteren Lichtblitzen, die so dargeboten wurden, daß ein geradliniger Bewegungseindruck entstand. Ein Fünftel der untersuchten MT-Neuronen zeigten im Vergleich zu der Stimulation mit einem einzelnen Lichtblitz eine starke Aktivierung, die mit der Anzahl der Lichtblitze zunahm. Mikami folgerte daraus einen starken inhibitorischen Mechanismus und einen schwachen fazilitatorischen Mechanismus, die für die Richtungsspezifität der MT-Neuronen verantwortlich sind.

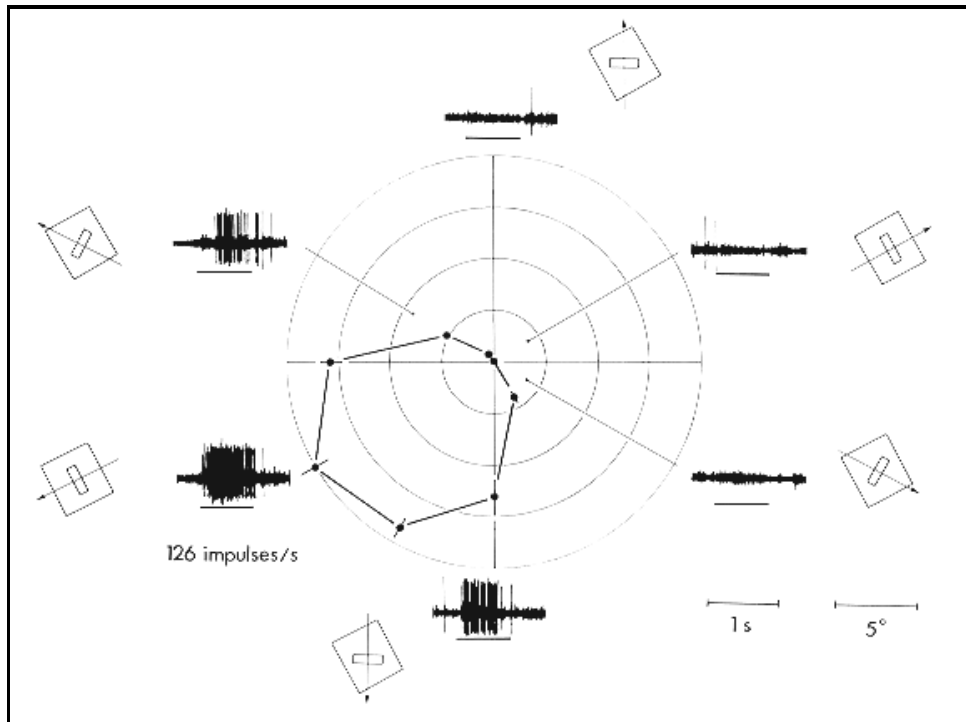


Abb. 2.16: Richtungsselektivität einer einzelnen Zelle aus MT. Die Aktivität eines MT-Neurons wurde in Abhängigkeit der Bewegungsrichtung des Stimulus gemessen. Die polare Darstellung der Aktivität ist ein Mittelwert aus fünf Wiederholungen bei zwölf Bewegungsrichtungen. Abbildung aus Maunsell und van Essen (1983).

Untersuchungen von Dinse und Krüger (1994) in verschiedenen Gehirnarealen der Katze zeigten eine große Standardabweichung bei der Verteilung der Latenzen in der Signalübertragung. So antworten Zellen des CGL nach dem Anschalten des Stimulus mit einer Verzögerung von 20 ms bis 50 ms. Dabei scheint die absolute Streuung im X-Pfad deutlich größer zu sein als die im Y-Pfad (Bowling 1989). Die Neuronen des PMLS zeigen einen maximalen zeitlichen Unterschied in ihrer Antwort von etwa 100 ms (Dinse und Krüger 1994).

Allman et al. (1985) zeigten, daß ein sich bewegendes Balkenreiz oder ein sich bewegendes Fleckenmuster innerhalb eines rezeptiven Feldes einer MT-Zelle durch das Hinzufügen eines sich in die gleiche Richtung bewegendes Umfeld die Antwort der Zelle in 44% der Fälle unterdrückt. Sie folgerten daraus einen Zentrum/Umfeld-Antagonismus der rezeptiven Feldstruktur zur Detektion von Kontrasten. Manche Zellen werden aber auch durch eine Bewegung im Umfeld in abweichenden Richtungen fasziliert. Diese Untersuchungen wurden zum Beispiel durch Tanaka et al. (1986) am Makaken und durch Hammond und MacKay (1981) sowie Hammond und Smith (1983) an der Katze bestätigt.

Die meisten MT-Neuronen sind zusätzlich sensitiv für binokulare Disparität (Bradley et al. 1995), so daß unter der Berücksichtigung von psychophysischen Effekten, wie zum Beispiel der *barberpole illusion* oder der Bewegungswahrnehmung transparenter Bewegungsfelder (z.B. Qian et al. 1994), von einer starken Interaktion zwischen der Verarbeitung von Tiefeninformation und Bewegungsinformation ausgegangen werden muß.

Rodman und Albright stießen 1989 bei der Untersuchung von MT-Neuronen mit „Barberplaid-Stimuli“ auf zwei Typen von Neuronen. Die Neuronen des ersten Typs waren sensitiv für die Bewegung des einzelnen Gitters, die des zweiten Typs ($\approx 30\%$) sensitiv für die Bewegung des gesamten Musters. Stoner und Albright stellten 1993 den direkten Zusammenhang zwischen Perzeption und neuronaler Aktivität her. Sie stellten fest, daß MT-Neuronen, die spezifisch auf das Gesamtmuster reagierten, bei Stimulation mit beiden Gittern, bei denen

allerdings Tiefenhinweise auftraten, so daß keine kohärente Bewegung eines einzelnen Gesamtmusters wahrgenommen werden konnte, zu Typ I- Neuronen wurden.

Lagae et al. (1993) fanden MT-Neuronen mit rezeptiven Feldern innerhalb einer Exzentrizität von 3°, die besonders selektiv für kleine Geschwindigkeiten waren. Diese scheinen aufgrund der Lage der rezeptiven Felder in der Fovea und der Sensitivität für kleine Geschwindigkeiten besonders gut zur Objektverfolgung geeignet zu sein.

Ein weiteres Areal, das hauptsächlich der Bewegungsverarbeitung dient, ist MST, das mediosuperiortemporale Areal. MST beinhaltet Zellen mit sehr großen rezeptiven Feldern. Die Zellen sind sensitiv für Rotation, Expansion und Kontraktion und bekommen ihre Eingabe von Neuronen in MT (z.B. Duffy und Wurtz 1991, Duffy und Wurtz 1995). MST könnte somit für die Bestimmung der Eigenbewegungsrichtung wichtig sein (z.B. Lappe 1996). Manche Neuronen in MT sind sensitiv für Geschwindigkeitsgradienten und stellen somit geeignete Grundeinheiten zur Analyse des optischen Flusses in MST dar (Treue und Andersen 1996).

2.3 Grundlagen der Modellierung

In diesem Kapitel werden zunächst die grundlegenden Modelle der Bewegungsdetektion beschrieben (Kapitel 2.3.1) und danach das in dieser Arbeit verwendete Modellneuron näher erläutert (Kapitel 2.3.2). Da das Modellneuron den Anforderungen der zur Verfügung stehenden Hardware genügen soll, wurde es für diese leicht modifiziert. Diese Modifikationen und die damit verbundenen Einschränkungen werden in Kapitel 2.3.3 dargestellt.

2.3.1 Methoden der Bewegungsdetektion

Zunächst werden einige einfache technische Methoden zur Bewegungsdetektion vorgestellt. Anschließend werden grundlegende Modelle zur Bewegungsdetektion beschrieben, die auf neuronalen Netzen basieren.

Das einfachste Verfahren, bewegte Objekte im Bild zu extrahieren, ist das Änderungsverfahren. Hierzu wird die Differenz zwischen zwei aufeinanderfolgenden Bildern ermittelt. Alle statischen Bereiche des Bildes führen zu einer minimalen Differenz und es bleiben nur sich bewegende Objekte übrig. Dieses Verfahren bietet sich allerdings nur für den einfachen Fall eines statischen Hintergrundes und gleichbleibender Beleuchtung an. Um Objektbewegungen sicher zu detektieren, können durch Rauschen verursachte Änderungen durch einen zusätzlichen Schwellwertmechanismus eliminiert werden. Bei dieser Methode findet keine Gruppierung zu einzelnen Objekten statt und es wird keine Bewegungsrichtung bestimmt.

Um die Einflüsse durch Beleuchtungswechsel zu verringern, kann das Änderungsverfahren auf signifikante räumliche Grauwertänderungen angewandt werden. Bei der Bewegungsdetektion mittels des sogenannten Gradientenverfahrens ergibt sich die Geschwindigkeit v als Quotient aus zeitlicher und räumlicher Änderung der Intensität I

$$v \cdot \nabla I = - \frac{\partial I}{\partial t}. \quad (\text{Gl 2.1})$$

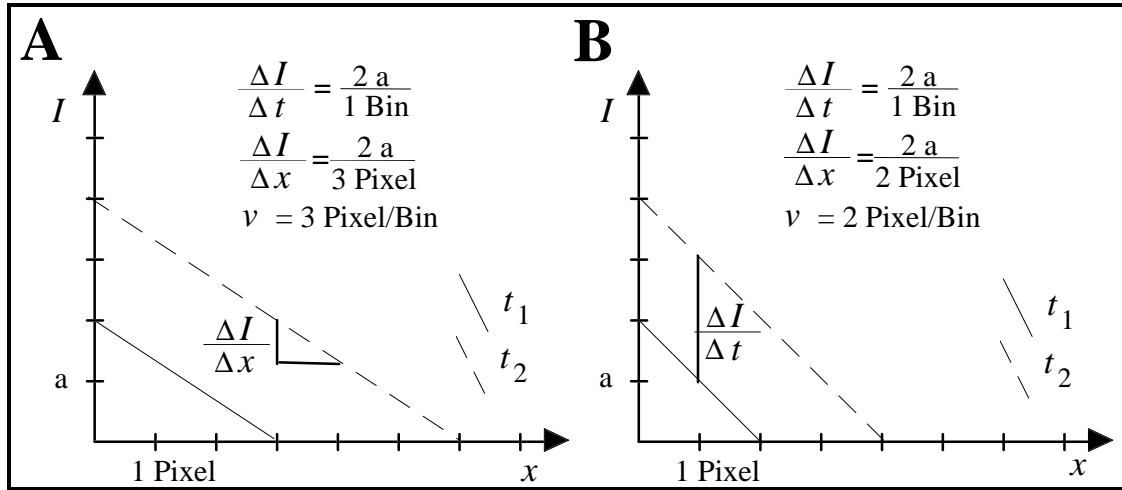


Abb. 2.17: Das Gradientenverfahren. Die Abbildung veranschaulicht das Gradientenverfahren anhand zweier Beispiele A und B. Zur Berechnung der Geschwindigkeit wird Gleichung 2.1 verwendet.

Das Gradientenverfahren wird anhand der Bewegung zweier eindimensionaler Kanten in Abbildung 2.17 dargestellt. Für eine Dimension wird Gleichung 2.1 zu

$$v_x = -\frac{\partial I / \partial t}{\partial I / \partial x} \quad \text{mit} \quad \frac{\partial I}{\partial t} = I(x, t_2) - I(x, t_1), \quad t_2 = t_1 + \Delta t, \quad \Delta t = 1 \text{ Bin},$$

$$\frac{\partial I}{\partial x} = I(x_2, t) - I(x_1, t), \quad x_2 = x_1 + \Delta x \quad \Delta x = 1 \text{ Pixel}.$$

Bei Kenntnis der Kantensteigung läßt sich durch Bestimmung der lokalen Intensitätsänderung die Geschwindigkeit v bestimmen. In Abbildung 2.17 sind zwei unterschiedlich steile Kanten mit verschiedenen Geschwindigkeiten dargestellt. Die Berechnung der Intensitätsänderung liefert für beide Kanten den gleichen Wert. Wird dieser Wert durch die Kantensteigung dividiert, so ergeben sich die beiden Kantengeschwindigkeiten. An Gleichung 2.1 wird auch ein Nachteil des Gradientenverfahrens deutlich. Die Geschwindigkeit kann nur orthogonal zur Kantenorientierung bestimmt werden, und auch nur dann, wenn die Funktion zur Beschreibung der Kante räumlich differenzierbar ist.

Da bei allen Verfahren die Bestimmung der Bewegung mit Fehlern behaftet oder wegen des Korrespondenzproblems das Ergebnis mehrdeutig ist, werden in der technischen Bildverarbeitung oft räumliche Glättungen über die lokalen Bewegungsdetektorantworten durchgeführt, wodurch ein Bewegungsfeld entsteht. Dabei wird gefordert, daß die lokalen Detektorantworten dem Bewegungsfeld ähnlich (Ähnlichkeitsbedingung) und daß Änderungen innerhalb des Bewegungsfeldes gering (Glattheitsbedingung) sind.

Bewegungsdetektoren, die auf dem Korrelationsverfahren (Reichardt 1957, Übersicht in Nagel 1985) beruhen, versuchen ein bestimmtes Muster, wie zum Beispiel eine Grauwertverteilung oder eine Anordnung von Intensitätskontrastkanten, in aufeinanderfolgenden Bildern wiederzufinden. Dazu wird die Ähnlichkeit einzelner Bildbereiche in zwei aufeinanderfolgenden Bildern innerhalb eines festzulegenden Suchbereiches über eine Kreuzkorrelation bestimmt. Die Bildbereiche mit der größten Ähnlichkeit legen somit die wahrscheinlichste Bewegungsrichtung fest. Große Bild- und Suchbereiche führen dabei zu hohen Rechenzeiten. Abhilfe schafft hierbei die Verwendung mehrerer Auflösungen, so daß zunächst eine grobe Bewegungsrichtung bestimmt wird und der Suchbereich für die feineren Auflösungen eingeschränkt werden kann.

Das technische Gradientenverfahren ist Grundlage vieler neuronaler Bewegungsdetektoren. Die Basis dieser Detektoren ist die multiplikative Verknüpfung der räumlichen Intensitätsänderung mit der zeitlichen. Diese Verschaltung geht auf Reichardt (1961) zurück (siehe auch Marr und Ullman (1981)). Wang et al. (1989) bestimmen den räumlichen Gradienten der Intensitätsverteilung durch eine Faltung mit einer Mexikanerhut-Funktion (eine zweifach differenzierte Gaußfunktion). Das Faltungsergebnis wird anschließend *sustained cells* und *transient cells* zugeführt. Der erste Zelltyp liefert dabei den räumlichen Gradienten mit Polarität der Kante, der zweite Zelltyp die zeitliche Änderung. Dies wird auch in Grossbergs Modell zur Bewegungsdetektion - dem *Motion Oriented Contrast (MOC)-Filter* - aufgenommen (Abbildung 2.18 A, Grossberg und Rudd 1992, Grossberg und Mingolla 1993, Francis und Grossberg 1996). Nach der Vorverarbeitung mittels Faltung mit einer Mexikanerhut-Funktion werden orientierte Kanten und zeitliche Transienten detektiert. Grossberg unterscheidet hierbei Transientenzellen, die auf zeitliche Intensitätserhöhung bzw. Intensitätsreduzierung reagieren. Durch die Kombination der Orientierungsdetektoren mit den beiden Transientendetektoren kann unabhängig von der Polarität der Kante ihre Bewegungsrichtung bestimmt werden. Zusätzlich findet noch eine räumliche Integration der Bewegungsinformation statt. Die Zellen der letzten Stufe stehen über inhibitorische Verbindungen in einem Wettbewerb, so daß das bewegte Objekt genauer lokalisiert werden kann. In einer Erweiterung des Modells wird zusätzlich noch die Bewegungsrichtung der Linienenden, die im Gegensatz zu der Bewegungsrichtung der Kanten eindeutig ist, genutzt. Die Linienenden werden durch Hyperkomplex-Zellen im *Boundary Contour System (BCS)* detektiert und dem MOC-Filter zugeführt. Dabei werden die Linienendenbewegungsrichtungen stärker gewichtet als die Bewegungsrichtungen der Kante. Eine weiterer Ausbau des Modells auf mehrere interagierende Auflösungen ermöglicht ferner die Bestimmung der Geschwindigkeit (Chey et al. 1998).

Das wohl bekannteste neuronale Modell zur Bewegungsdetektion ist der bereits Ende der sechziger Jahre von Reichardt (Reichardt 1957, Reichardt 1961, Egelhaaf und Reichardt 1987, Egelhaaf et al. 1988, Poggio und Reichardt 1973) vorgestellte Bewegungsdetektor. Der nach ihm benannte *Reichardt-Detektor* (Abbildung 2.18 B) besteht aus zwei Rezeptoren im Abstand s , deren Ausgänge jeweils um die Dauer Δ verzögert und unverzögert an zwei Multiplizierer gekoppelt sind. Die Differenz der Produkte ergibt das Membranpotential. Dieser Detektor ist ein Kovarianzdetektor. Wird zunächst der Rezeptor R1 und danach der Rezeptor R2 mit einer Verzögerung von $t = \Delta$ durch eine Objektbewegung erregt, gelangen die beiden Aktivitäten gleichzeitig zu Multiplizierer M1, was zu einem positiven Membranpotential - oder nach Vergleich mit einer statischen Schwelle zu einer Eins - des Detektors führt. Der Multiplizierer M2 gibt hingegen eine Null aus. Somit detektiert der Detektor eine Geschwindigkeit von $v = s / \Delta$. Diese Geschwindigkeit würde er aber auch bei der ständigen Reizung beider Rezeptoren, wie sie durch ein stehendes Objekt verursacht werden könnte, detektieren. Deswegen werden die beiden Ausgänge der Multiplizierer wechselseitig voneinander abgezogen. Damit wird keine Bewegung detektiert, wenn beide Multiplizierer aktiv sind. Ein Reichardt-Detektor kann deswegen auf zwei entgegengesetzte Bewegungsrichtungen antworten.

Äquivalent zum Reichardt-Detektor kann die Bewegungsdetektion auch im Frequenzraum erfolgen (Adelson und Bergen 1985, Watson und Ahumada 1985). Dazu können zum Beispiel Gabor-Funktionen, die aus einer Multiplikation einer Gauß-Funktion mit einer Sinus- oder Cosinus-Funktion entstehen, zur Faltung mit dem Bild benutzt werden (z.B. Grzywacz und Yuille 1991). Ein Reichardt-Detektor, der seine Eingabe über raumzeitliche Filter erhält, wird mit *elaborated Reichardt detector* bezeichnet (van Santen und Sperling 1984).

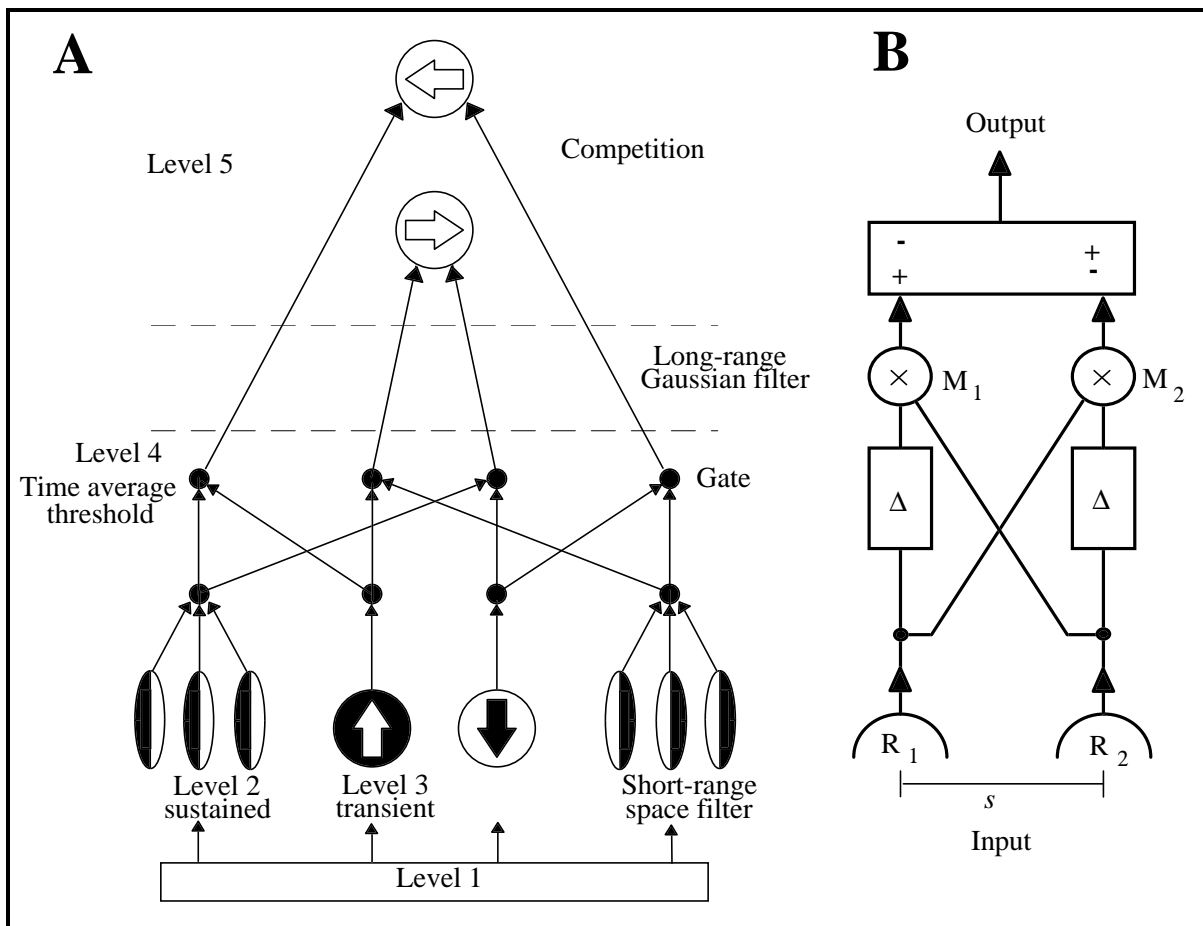


Abb. 2.18: Modelle zur Bewegungsdetektion. A: Grossbergs MOC-Filter nach Francis und Grossberg (1996). B: Der Reichardt-Detektor (Reichardt 1961). Nähere Erläuterungen im Text.

Der von Barlow und Levick (1965) vorgestellte Bewegungsdetektor beruht auf neurophysiologischen Untersuchungen an der Kaninchenretina (Barlow und Hill 1963, Barlow et al. 1964). Um die Richtungsselektivität retinaler Ganglienzellen des Kaninchens zu erklären, schlugen sie unidirektionale laterale Verbindungen zwischen den Zellen vor, die in Vorzugsrichtung exzitatorisch und in Gegenrichtung inhibitorisch wirken. Findet die Bewegung nun in Vorzugsrichtung statt, so erfolgt durch die exzitatorische Kopplung eine Fazilitation, wohingegen eine Bewegung in die nicht bevorzugte Richtung eine Aktivitätsreduzierung zur Folge hat. Diese Verschaltung führt damit zu einer Bewegungsrichtungsdetektion.

Die oben beschriebenen Richtungsdetektoren unterliegen dem Apertur- und dem Korrespondenzproblem (Übersicht in Nakayama (1985)), d.h., das Ergebnis der lokalen Bewegungsanalyse ist mehrdeutig. Um eine eindeutige Bewegungsrichtung zu erhalten, sind die drei in Abbildung 2.19 dargestellten Algorithmen denkbar (Groh et al. 1997). Einerseits besteht die Möglichkeit, die Bewegungsdetektoren, die einen bestimmten Ort des Bildes verarbeiten, als Population auszuwerten (*Populationskodierung*), andererseits können die mehrdeutigen Antworten dieser Neuronengruppe auf eine, z.B. die stärkste Antwort, reduziert werden (*winner take all*). Die Reduzierung auf die stärkste Antwort wird zum Beispiel von Glünder und Lehmann (1992) durch einen Relaxationsprozeß bewältigt, indem die Antworten der räumlich benachbarten Detektoren miteinbezogen werden.

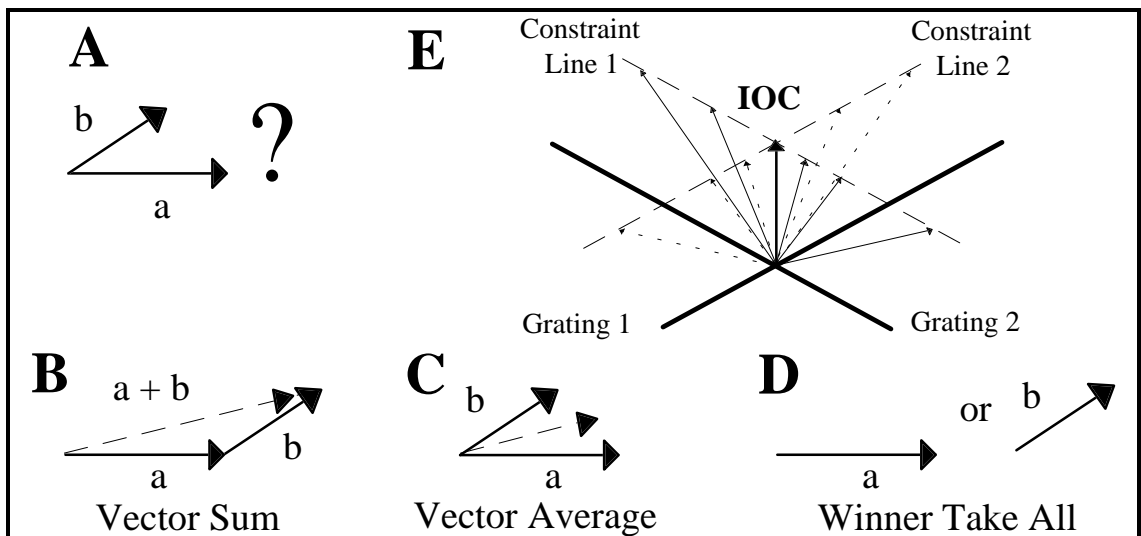


Abb. 2.19: Methoden zur Generierung einer eindeutigen Bewegungsrichtungsantwort. A: Zwei Bewegungsrichtungen an einem Ort, die zu einer über Vektor-Sumation (B), Vektor-Mittelung (C) oder *winner take all* (D) zu einer einzigen Bewegungsrichtungsantwort werden. E: *Intersection of Constraints*.

Diese Verfahren können auch bei der Integration verschiedener räumlich benachbarter Bewegungsdetektorantworten, wie sie z.B. bei der *barberplaid illusion* auftreten, verwendet werden. Ein weiterer Ansatz, der oft zur Erklärung der *barberplaid illusion* herangezogen wird, ermittelt übereinstimmende Richtungsdetektorantworten (*Intersection of Constraints* (IOC), Adelson und Movshon 1982). In der Abbildung 2.19 E sind zwei Linien und ihre möglichen Bewegungsrichtungen dargestellt. Nur eine Bewegungsrichtung tritt bei beiden Linien auf. Sie spiegelt die wahrgenommene Bewegungsrichtung der beiden überlagerten Linien wider. Alle möglichen Bewegungsvektoren bilden eine Begrenzungslinie. Der Schnittpunkt der beiden Begrenzungslinien legt die Bewegungsrichtung der beiden Linien fest.

2.3.2 Das Marburger Modellneuron

Bei der Simulation neuronaler Netze sollten die Eigenschaften des Modellneurons mit dem Abstraktionsgrad der Untersuchung korrelieren. Gemäß der zugrundeliegenden Marburger Arbeitshypothese, daß Objekte durch synchrone neuronale Aktivität repräsentiert werden (Reitboeck 1980 und 1983 sowie von der Malsburg 1981), muß das Modellneuron zu definierten Zeitpunkten Aktionspotentiale generieren können, was für Synchronisationsuntersuchungen zwingend notwendig ist. Dies setzt wie beim biologischen Neuron einen Schwellenmechanismus und einen Mechanismus zur Simulation der Refraktärzeit voraus. Ferner muß zusätzlich zu den zum Aufbau von rezeptiven Feldern erforderlichen Verbindungen - den sogenannten *Feeding*-Verbindungen - eine Kopplung zwischen den Modellneuronen existieren, die eine Synchronisation ihrer Ausgangsaktivität ermöglicht.

Das Marburger Modellneuron (Eckhorn et al. 1990) bietet diese Eigenschaften. Es basiert auf dem 1970 von French und Stein entwickelten *neural analog*, wurde aber durch ein auf das *Feeding*-Potential $F(t)$ modulatorisch wirkendes *Linking*-Potential $L(t)$ erweitert. Das Membranpotential $U(t)$ eines Modellneurons mit K Dendriten zum Zeitpunkt t berechnet sich dann nach Gleichung 2.2.

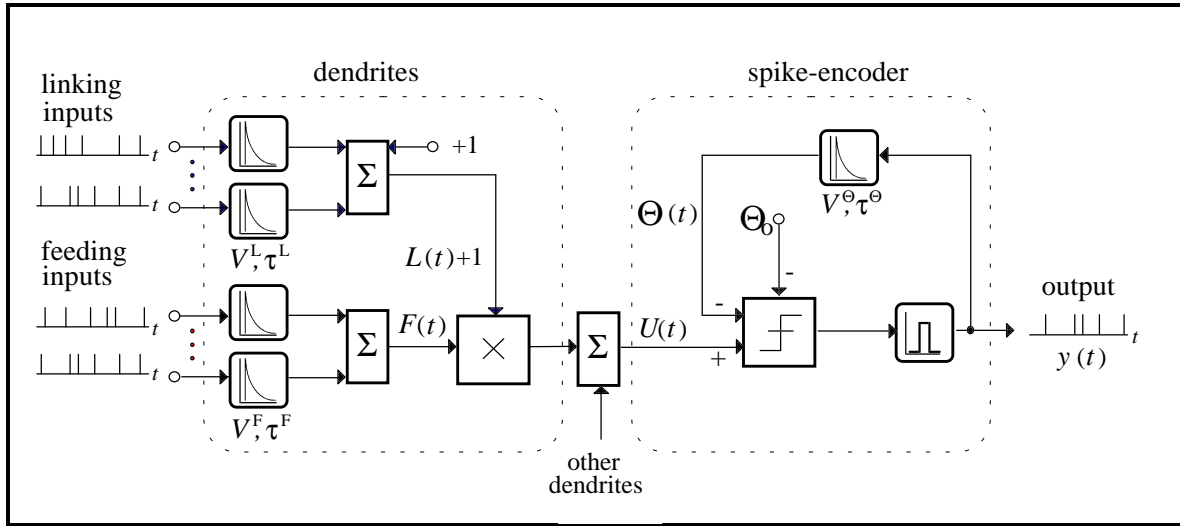


Abb. 2.20: Schematischer Aufbau des Marburger Modellneurons. Das Marburger Modellneuron besteht aus einem dendritischen und einem somatischen Teil (Spike-Encoder). Die Eingangsimpulse werden durch Leckintegratoren, charakterisiert durch die Parameter V und τ , zeitlich integriert. Das Membranpotential $U(t)$ ergibt sich aus der Multiplikation der postsynaptischen Potentiale $F(t)$ und $L(t)+1$. Übersteigt $U(t)$ die Summe aus dynamischer Schwelle $\Theta(t)$ und Schwellenoffset Θ_0 , so wird am Ausgang des Neurons ein Aktionspotential generiert und $\Theta(t)$ über einen Leckintegrator inkrementiert. (Modifiziert nach Eckhorn et al. 1990.)

$$U(t) = \sum_{k=1}^K F_k(t) \cdot (1 + L_k(t)). \quad (\text{Gl. 2.2})$$

Das Potential eines Dendriten zum Zeitpunkt t ist somit nur von Null verschieden, wenn auch das Feeding-Potential $F(t)$ ungleich Null ist: Das Linking-Potential $L(t)$ kann das Modellneuron allein nicht erregen.

Durch die Verwendung von Leckintegratoren an den Eingängen des Modellneurons erfolgt eine zeitliche Integration der Eingangsimpulse (von Foerster 1966) entsprechend dem postsynaptischen Potential des biologischen Neurons. Damit ergibt sich das postsynaptische Potential $P(t)$ aus der Faltung der Eingangsimpulse $x(t)$ mit der Stoßantwort eines Tiefpasses $h(t)$ zu

$$P(t) = x(t) * h(t), \quad (\text{Gl. 2.3})$$

mit

$$h(t) = \begin{cases} V \cdot \exp(-t / \tau), & \text{für } t \geq 0 \\ 0, & \text{für } t < 0 \end{cases}. \quad (\text{Gl. 2.4})$$

Die Gleichung 2.3 gilt für die Berechnung des Feeding-Potentials $F(t)$ und des Linking-Potentials $L(t)$. Die Parameter V und τ aus Gleichung 2.4, die die Stoßantwort $h(t)$ eines Tiefpasses charakterisieren, sind in der Regel für $F(t)$ und $L(t)$ unterschiedlich zu wählen.

Das Marburger Modellneuron generiert in Anlehnung an das biologische Neuron ein Aktionspotential, wenn das Membranpotential $U(t)$ ein bestimmtes Schwellenpotential Θ_0 übersteigt (Gl. 2.5). Die Refraktärzeit biologischer Neuronen wird nachgebildet, indem zusätzlich ein dynamisches Schwellenpotential $\Theta(t)$ implementiert wird, das bei der Erzeugung eines Aktionspotentials um den Wert V^Θ inkrementiert und über einen weiteren Leckintegrator dekrementiert wird (Gl. 2.6). Mit der dynamischen Schwelle kann somit die relative Refraktärzeit - die Hyperpolarisation nach der Generierung eines Aktionspotentials - simuliert

werden. Die absolute Refraktärzeit wird durch die Zeitdiskretisierung bei digitaler Hardware modelliert. In einem diskreten Simulationsschritt (1 Bin) kann maximal ein Aktionspotential generiert werden, womit dieser einer Millisekunde entspricht.

$$y(t) = \begin{cases} 1, & \text{für } U(t) \geq (\Theta(t) + \Theta_0) \\ 0, & \text{sonst} \end{cases} \quad (\text{Gl. 2.5})$$

$$\Theta(t) = y(t) * h^\ominus(t) \quad (\text{Gl. 2.6})$$

Mit $y(t)$ wird das Ausgangssignal des Modellneurons zur Zeit t bezeichnet. $h^\ominus(t)$ steht für die Stoßantwort eines Tiefpasses (Gl. 2.4) mit den Parametern V^\ominus und τ^\ominus .

2.3.3 Der Hardware Accelerator

2.3.3.1 Motivation und Simulationsumgebung

Das im nachfolgenden Kapitel 3 dargestellte technische System zur visuellen Informationsverarbeitung ist hierarchisch aufgebaut. In jeder Verarbeitungsstufe werden Merkmale extrahiert, was zu einer geringen Anzahl an aktiven Modellneuronen - einer spärlichen Kodierung - führt. Für eine algorithmische Implementation des Marburger Modellneurons sollte somit primär die Anzahl an generierten Aktionspotentialen bestimmend für die Rechenzeit sein. Dies kann durch eine sender-orientierte Verarbeitungsstruktur erreicht werden. Das von Frank und Hartmann (1995) propagierte Konzept eines „Neuronale Netze Beschleunigers für impulsodierende Modellneuronen“, kurz *Accelerator* genannt, bietet diese Eigenschaft. Eine von Jahnke (1995) durchgeführte Untersuchung (Tabelle 2.1) zeigt, daß sowohl der hier verwendete Accelerator als auch sein Nachfolger *NESPINN* der Echtzeitanforderung technischer Systeme gerecht werden.

Die Verarbeitung der Bilddaten, die sowohl direkt von einer Kamera als auch aus einem Bilddatenspeicher geliefert werden können, erfolgt in mehreren Stufen (Abb. 2.21). Die erste Verarbeitungsstufe wird von der Vorverarbeitung, die auf einem schnellen Computer oder einer speziellen Hardware ausgeführt wird, gebildet. Eine direkte Speisung des Hardware-Accelerators oder des NESPINN mit den Bilddaten ist aufgrund der zu großen Datenmenge und der begrenzten Datenübertragungsrate nicht möglich. Folglich hat eine erste Informationsreduktion durch die Vorverarbeitung zu erfolgen. In Anlehnung an neurophysiologische Untersuchungen, siehe hierzu auch Kapitel 2.2.2, werden in den Bilddaten Intensitätskontraste extrahiert und diese impulsodiert als Adressen aktiver Modellneuronen an die nächste Verarbeitungsstufe übergeben. Die Vorverarbeitung wird in Kapitel 3.1 genauer beschrieben.

Modellneuronen -anzahl	SPAR C 20	CNAPS	TI-C80	CM-2	Accelerator	NESPINN
16K	40 ms	1,5 ms	~ 2 ms	< 0,01 ms	0,2 ms	< 0,1 ms
128K	500 ms	-	~ 400 ms	< 0,1 ms	1,5 ms	< 1 ms
512K	> 2 s	-	> 1 s	~ 1 ms	~ 2 ms	~ 1 ms

Tab. 2.1: Benötigte Rechenzeit verschiedener Hardware für einen Simulationsschritt (1 ms) bei einer Netzwerkaktivität von 0,01%. Die Zeiten für Sparc 20 und CNAPS wurden gemessen, die anderen geschätzt (Jahnke 1995).

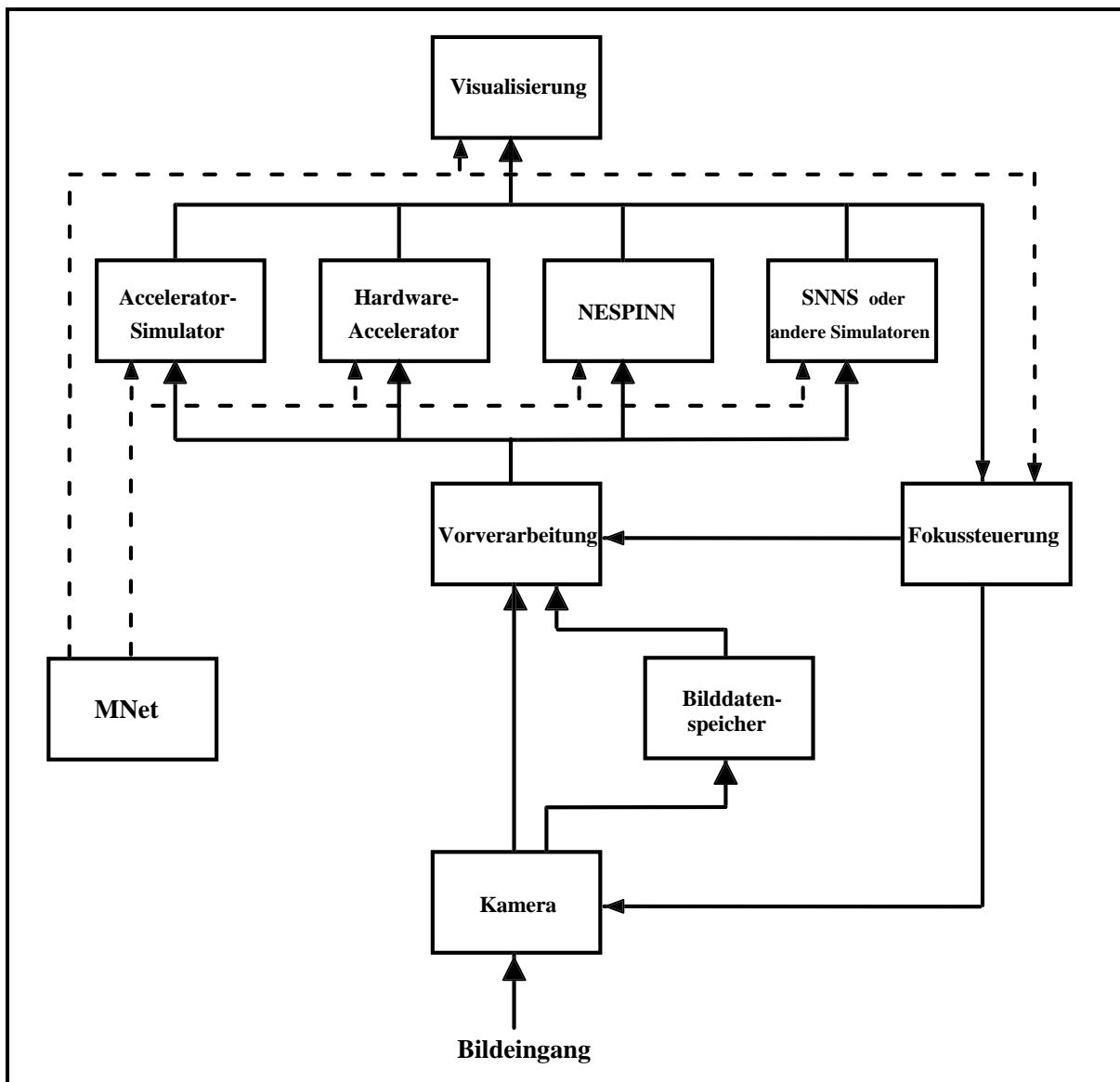


Abb. 2.21: Simulationsumgebung. Dargestellt ist der Informationsfluß (durchgezogene Linie) von der Eingabe bis zur Visualisierung des Simulationsergebnisses sowie die Verknüpfung des MNet-Compilers mit den verschiedenen Simulatoren und dem Visualisierungsmodul. Die Kamera erzeugt Bilder, die entweder direkt oder über einen Bilddatenspeicher auf die Vorverarbeitung gegeben werden. Die von der Vorverarbeitung erzeugten Daten werden dann von einem der Simulatoren weiterverarbeitet. Anschließend werden die Simulationsergebnisse an das Visualisierungsmodul oder die Fokussteuerung zur Auswertung übergeben. Das MNet-Modul dient zur Übertragung der Netzwerkarchitektur auf die Simulatoren und stellt die Topographie der Modellneuronen zur Verfügung.

Für das neuronale Netz stehen in der nachfolgenden Verarbeitungsstufe mehrere Plattformen, entweder als Hardware oder als Software-Simulator, zur Verfügung. Die Hardware bietet Vorteile in der Rechenzeit, ist aber aufgrund der begrenzten Übertragung an Ausgabedaten nachteilig beim Konstruieren und Optimieren Neuronaler Netze. Ein weiterer wichtiger Grund für die Verwendung von Software-Simulatoren ist die Überprüfung der Leistungsfähigkeit der in Hardware zu realisierenden Modellneuronen. Deswegen wurden Funktionalität und Rechengenauigkeit des Accelerator-Neurons in einem Accelerator-Simulator (Spengler 1995, Gabriel 1995) nachgebildet. Mit Ausnahme der Vorverarbeitung wurden alle nachfolgenden Simulationen mit diesem Accelerator-Simulator durchgeführt. Andere Simulatoren, wie zum Beispiel der Stuttgarter Neuronale Netze Simulator (SNNS), können per se auch eingesetzt werden, sind aber in der Regel nicht für die Simulation impulskodierender Modellneuronen konzipiert worden.

Bei dieser Vielfalt an Simulationsmöglichkeiten ist es sinnvoll, ein von der Plattform unabhängiges neuronales Netz erzeugen zu können. Diese Möglichkeit bietet MNet, die Marburger Netzwerkbeschreibungssprache (Möller 1995). MNet wurde zur einfachen Beschreibung von komplexen Netztopologien entwickelt. Im Gegensatz zu konventionellen Sprachen der dritten Generation lassen sich große Netzwerke mit Hilfe mächtiger MNet-Befehle leicht modularisieren, was unabdingbar für eine arbeitsteilige Entwicklung großer Systeme ist.

Die Initialisierung der Hardware erfolgt über mehrere von MNet erzeugten Dateien, die auch optional für die Software-Simulatoren verwendet werden können. Die Formate der Dateien für unterschiedliche Hardware differieren hingegen voneinander. Zur Auswertung ist eine anschließende Visualisierung der Ergebnisse notwendig. Da die Simulatoren nicht über die Topographie der Modellneuronen verfügen, ist diese ebenfalls von MNet bereitzustellen.

Die Simulationsergebnisse können aber auch sofort weiterverarbeitet werden. Aktive Sehsysteme sollten in der Lage sein, Objekte verfolgen und in Abhängigkeit zur Aufmerksamkeitssteuerung neue Objekte fokussieren zu können. Dies wird durch eine von Pauly et al. (1997) entwickelte Fokussteuerung erreicht.

2.3.3.2 Zeitschritt-Simulationsverfahren

Wie in Kapitel 2.3.2 dargestellt wurde, setzt sich das Marburger Modellneuron aus einem dendritischen und einem somatischen Bestandteil zusammen. Das Modellneuron wird über die Dendriten erregt und generiert nach einem Vergleich des resultierenden Membranpotentials mit der Schwelle sein Ausgangssignal. Diese Unterteilung in eine Erregungsphase und in eine Phase, in der zum einen das Membranpotential berechnet wird, zum anderen aber auch die postsynaptischen Potentiale sowie die dynamische Schwelle dekrementiert werden müssen, wird auch in der Hardwarerealisierung des Accelerators beibehalten. Somit bilden Erregungsphase und die sogenannte Abklingphase die Basis des Zeitschritt-Simulationsverfahrens. Innerhalb eines Zeitschrittes werden eine Erregungs- und eine Abklingphase durchlaufen. Parallel zur Abklingphase ist eine externe Kommunikation sowie die Ausführung eines Lernverfahrens möglich. In der Abbildung 2.22 ist das Zeitschritt-Simulationsverfahren graphisch dargestellt.

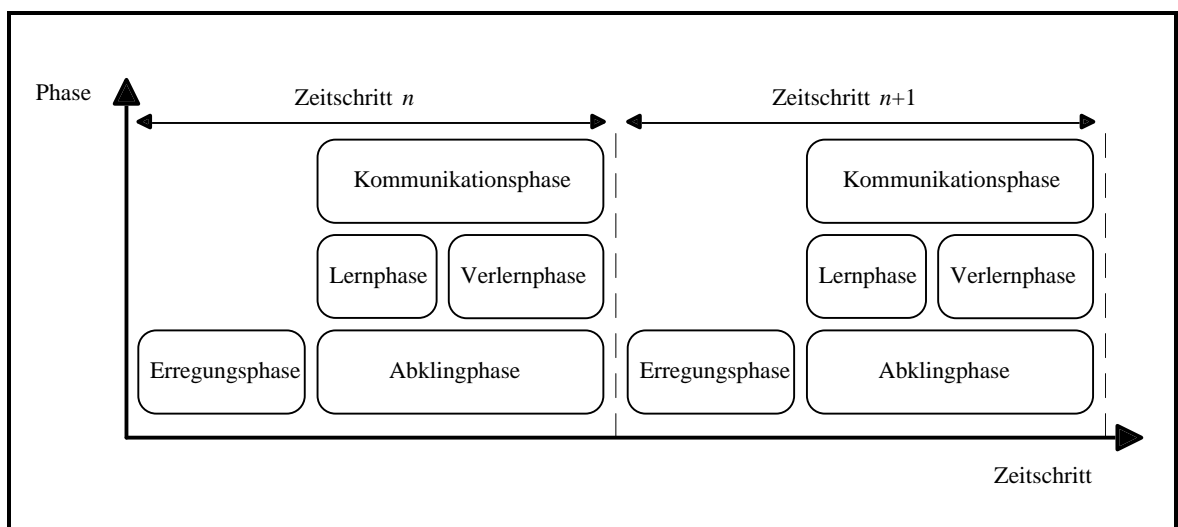


Abb. 2.22: Phasen der Zeitschritt-Simulation. Modifiziert nach Frank (1997).

2.3.3.3 Module des Hardware-Accelerators

Der Accelerator verfügt über eine sich funktional an die Phasen der Zeitschritt-Simulation anlehende modulare Struktur. Hierzu gehören die Module zur Kommunikation und Initialisierung sowie die Verbindungs-, Neuronen- und Lernmodule.

Das Kommunikationsmodul dient sowohl der Vernetzung mehrerer Acceleratorboards als auch der Anbindung an das Host-System, wodurch die Übertragung der Vorverarbeitungsdaten sowie die Beobachtung der Netzaktivität ermöglicht wird. Die externe Kommunikation erfolgt wie die interne ausschließlich über Listen mit Adressen aktiver Modellneuronen.

Mit Hilfe des Initialisierungsmoduls werden zu Beginn der Simulation die vom MNet-Compiler erzeugten Daten in den Speicher des Accelerators geladen. Der Speicher des Verbindungsmoduls nimmt hierbei die Netztopologie und die Gewichte auf. Das Neuronenmodul besteht aus dem Speicher für die Modellneuronenvariablen, dazu gehören insbesondere die Teilpotentiale, der Steuerungseinheit für die Abklingphase und der Einheit zur Berechnung des Membranpotentials. Der modulare Aufbau des Accelerators ermöglicht die Verwendung unterschiedlicher Lernregeln durch die optionale Verwendung verschiedener Lernmodule. Die Lernverfahren werden sich hierbei allerdings auf den Einsatz von Hebbschen Lernregeln beschränken.

Die Abbildung 2.23 zeigt schematisch den Informationsfluß zwischen Neuronen-, Verbindungs- und Lernmodul. Je nach Speicherbestückung können auf einem Acceleratorboard bis zu 128 K Modellneuronen simuliert werden. Dazu wird jedem Modellneuron eindeutig eine Nummer zugewiesen. Die Generierung eines Aktionspotentials bewirkt die Aufnahme der entsprechenden Modellneuronennummer in die Liste aktiver Modellneuronen. Diese Liste dient in der nachfolgenden Erregungsphase dem Verbindungsmodul zur Erregung der Empfängerneurone.

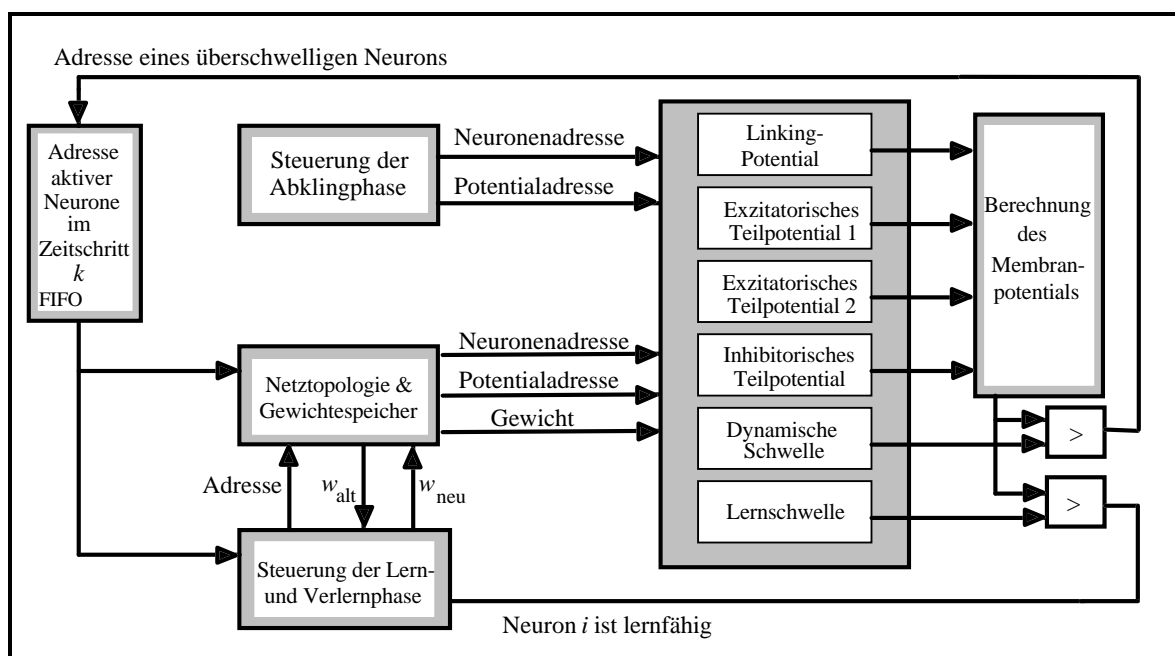


Abb. 2.23: Hardwarestruktur des Accelerators. Dargestellt sind die wichtigsten Bestandteile des Neuronenmoduls (Steuerung der Abklingphase, Teilpotentialspeicher, Einheit zur Berechnung des Membranpotentials), des Kommunikationsmoduls (FIFO-Liste), des Verbindungsmoduls (Netztopologie und Gewichtespeicher) und des Lernmoduls (Steuerung der Lern- und Verlernphase). Modifiziert nach Frank (1997).

In der Abklingphase werden dann die Berechnung der Membranpotentiale sowie das Abklingen der Teilpotentiale über Abklingtabellen durchgeführt. Die Verwendung von Abklingtabellen reduziert im Gegensatz zum Dekrementieren durch Multiplizieren zum einen den Aufwand an Hardware und Rechenzeit, ermöglicht aber auch zum anderen einen flexibleren Abklingverlauf, der nur der Einschränkung unterliegt, durch eine streng monotone Funktion beschreibbar sein zu müssen. Weiterhin wurde zum Minimieren der Abklingzeit eine Abklingliste implementiert. Nur Modellneuronen, bei denen mindestens ein Teilpotential über einem festlegbaren Potentialminimum liegt oder die in der Erregungsphase angesprochen wurden, werden in die Abklingphase einbezogen. Die zweite Bedingung ist zwingend notwendig, weil einmal unter das Potentialminimum abgefallene Teilpotentiale auch bei Erregung der entsprechenden Modellneuronen nicht mehr in die Abklingliste aufgenommen würden, da die Membranpotentialberechnung nur bei in der Abklingliste aufgeführten Modellneuronen durchgeführt wird.

2.3.3.4 Netztopologie und Gewichtespeicher

Wie bereits oben erwähnt wurde, dient das Verbindungsmodul zum einen zur Speicherung der Netztopologie und der Gewichte, zum anderen zur Erregung der Teilpotentiale. Hierzu wird eine senderorientierte verkettete Speicherung benutzt (Abbildung 2.24). Speicherbedingt stehen 16 M Verbindungen zur Verfügung. Um eine möglichst große Flexibilität zu gewährleisten, sollten die Verbindungen den Modellneuronen beliebig zugeordnet werden können. Dies könnte durch eine $N_S \times N_E$ -Gewichtsmatrix ermöglicht werden, wobei N_S und N_E der Anzahl der möglichen Senderneuronen respektive Empfängerneuronen entsprechen, was aber zu einer sehr ineffektiven Speicherausnutzung führen würde. Aus diesem Grunde ist der Gewichtespeicher in Erregungsinformationsblöcke (EIB) variabler Größe unterteilt, wobei jeder EIB Sektionen von modifizierbaren und nicht-modifizierbaren Gewichten enthält. Jede Zeile des Gewichtespeichers beinhaltet neben Empfängernummer, Teilpotential und Gewicht auch Informationen zur Steuerung, so daß durch die Nummer eines aktiven Senderneurons in Verbindung mit dem Blockstartadressespeicher (BSS) die der Topologie entsprechenden Empfängerneuronen erregt werden. Die Dauer der Erregungsphase hängt somit von der Anzahl aktiver Modellneuronen ab.

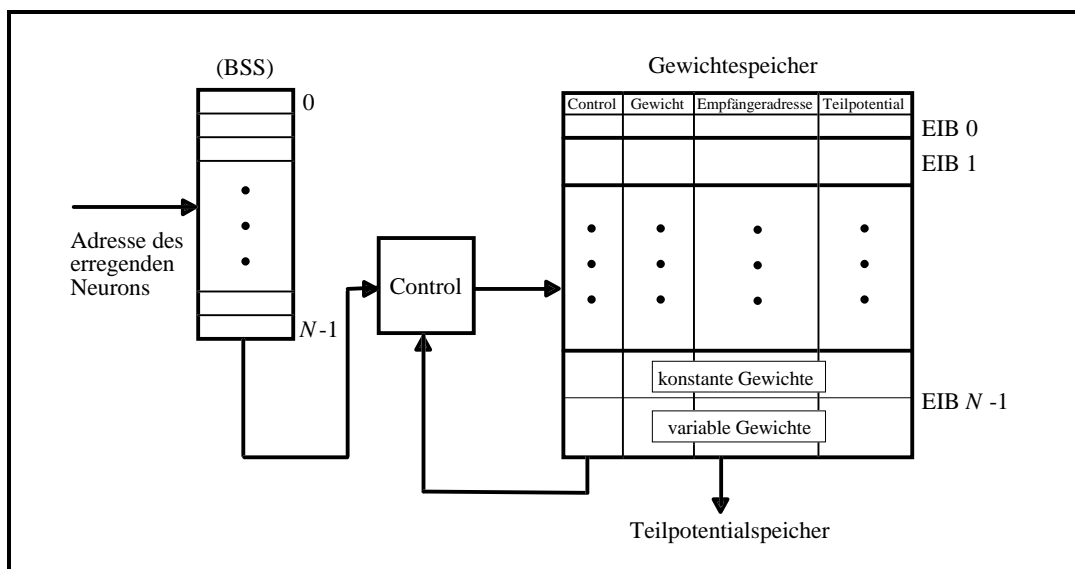


Abb. 2.24: Senderorientierte verkettete Speicherung der Netztopologie und der Gewichte. Der Blockstartadressespeicher (BSS) verweist auf den Beginn eines Erregungsinformationsblockes (EIB), wodurch in Verbindung mit einer Kontrolleinheit die Teilpotentiale aller Empfängerneuronen eines Senderneurons erregt werden können. Modifiziert nach Frank 1997.

2.3.3.5 Berechnung des Membranpotentials

Das Ausgangssignal eines Modellneurons wird aus postsynaptischen Potentialen und einer dynamischen Schwelle berechnet. Wie bereits in Kapitel 2.3.2 geschildert, verfügt das Marburger Modellneuron über zwei funktional verschiedene Eingänge. Somit sind mindestens drei Variablen pro Modellneuron zu speichern. Da jeder Variablen aber nur eine Abklingfunktion zugeordnet werden kann, sind bereits grundlegende Simulationen, wie sie in Arndt (1993), Dicke (1992), Stoecker (1993) und Schott (1994) beschrieben werden, nicht mehr durchführbar. Die Mindestanforderung an einen rückgekoppelten Assoziativspeicher sind ein Dendrit für die Aufnahme des Eingabemusters, ein Dendrit für Rückkopplungen sowie ein weiterer Dendrit für die Inhibition. Da alle Dendriten unterschiedliche Abklingverhalten besitzen, muß somit die Mindestanforderung an ein Modellneuron auf einen Linking-Eingang, zwei exzitatorische und einen inhibitorischen Feeding-Eingang sowie einer dynamischen Schwelle erhöht werden. Weitere Teilpotentiale wären zwar nützlich und könnten die Simulationsbedingungen erheblich vereinfachen, würden aber zu einer unerwünschten Erhöhung der Rechenzeit führen.

Im Vergleich zu Simulationen auf konventionellen Rechnern sind weitere Einschränkungen bei der Verwendung des Accelerators in Kauf zu nehmen. Die Berechnung des Membranpotentials erfolgt auf der Grundlage von 16-Bit-Operationen, wodurch die Amplitudenauflösung des Membranpotentials auf 16 Bit beschränkt wird. Dieser maximalen Wortbreite entsprechend unterliegen die einzelnen Teilpotentiale ebenfalls einer Einschränkung. So werden Linking- und Feeding-Potentiale in einer vorzeichenbehafteten Fixpunktdarstellung von s6.4, d.h. sechs Vorkommastellen, 4 Nachkommastellen sowie ein zusätzliches Vorzeichenbit, bzw. s9.5 repräsentiert. Die Nachkommastellen sind notwendig, damit in der Abklingphase die Teilpotentiale in kleinen Schritten dekrementiert und in der Erregungsphase auch sehr kleine Gewichte oder Gewichtsänderungen merkbar aufgenommen werden können. Dementsprechend wurden auch die Auflösungen der Gewichte gewählt: Linking-Gewichte werden in s4.4, Feeding-Gewichte in s3.5 Fix-Punkt-Darstellung angegeben. Wie in Schott (1994) gezeigt, ist für eine Objektseparation Membranpotential-Rauschen sehr wichtig. Da der Accelerator eine Rauschgenerierung nicht vorsieht, muß auf eine Inhomogenisierung der Gewichte durch kleine Gewichtsveränderungen zurückgegriffen werden, was denselben Effekt erzielt. Ein weiterer wichtiger Grund für die Verwendung von Nachkommastellen ist die Notwendigkeit, kleine Lernschritte, insbesondere bei Schwächung eines Gewichtes bei Hebbschem Lernen, ausführen können zu müssen. Des weiteren soll die Verschaltung vieler Modellneuronen aufeinander möglich sein, ohne daß die Teilpotentiale in die Sättigung laufen.

Die Abbildung 2.25 zeigt die einzelnen Teilpotentiale und ihre Verschaltung. Die beiden exzitatorischen Teilpotentiale werden zunächst addiert. Danach werden wie bei den beiden anderen Teilpotentialen, dem Linking-Potential und dem inhibitorischen Potential, nur positive Potentiale weiterverarbeitet. Ohne diese Einschränkung bestünde die Gefahr, daß ein negatives Linking-Potential und ein negatives Feeding-Potential ein großes, positives Membranpotential ergeben könnten. Damit das Linking-Potential ausschließlich modulatorisch auf das Feeding-Potential wirkt, muß es gemäß Gleichung 2.2 ein Minimum von $LP_{min}=1$ aufweisen, was durch die Abklingtabelle einfach zu realisieren ist. Ein Minimum von $LP_{min}=0$ würde zu einer Art AND-Gatter führen. Nur wenn sowohl Feeding- als auch Linking-Potential von Null verschieden sind, besteht die Möglichkeit der Aktionspotentialgenerierung.

Der Schwellenoffset muß ebenfalls über die Abklingtabelle in die dynamische Schwelle integriert werden. Ferner ist zu beachten, daß der Wert des Speichers der dynamischen Schwelle um ein Bit nach links verschoben wird. Dies ist erforderlich, da der maximale Wert der dynamischen Schwelle größer als der maximale Wert des Membranpotentials sein muß.

Wäre dies nicht der Fall, so würden beim Erreichen der Sättigung des Membranpotentials fortwährend Aktionspotentiale generiert werden. Ist das höchstwertige Bit der dynamischen Schwelle gesetzt, so kann auf keinen Fall ein Aktionspotential ausgelöst werden. Gleiches gilt für einen eventuellen Überlauf bei der Addition von dynamischer Schwelle und inhibitorischem Teilpotential. Das Ergebnis dieser Addition wird dann mit dem Membranpotential verglichen. Ist das Membranpotential größer als die Summe aus dynamischer Schwelle und inhibitorischem Potential und ist weder ein Überlauf bei der Addition aufgetreten noch das höchstwertige Bit der dynamischen Schwelle gesetzt, dann wird ein Aktionspotential generiert, d.h., das Neuron wird in die Liste aktiver Modellneuronen aufgenommen. Für die dynamische Schwelle gelten noch zwei weitere Besonderheiten: Das Schwelleninkrement DS_{Ink} kann nur 16 oder ein Vielfaches davon sein, da die vier niederwertigen Bits nicht gesetzt werden können. Die zweite Einschränkung stellt die geringe Anzahl an möglichen Abklingtabellen dar. Während für das Linking-Potential 32 und für die Feeding-Potentiale jeweils 4 Abklingtabellen pro Accelerator zur Verfügung stehen, sind nur zwei Abklingtabellen für die dynamische Schwelle verfügbar. Dies ist in der internen Speicherung der Abklingtabellen begründet.

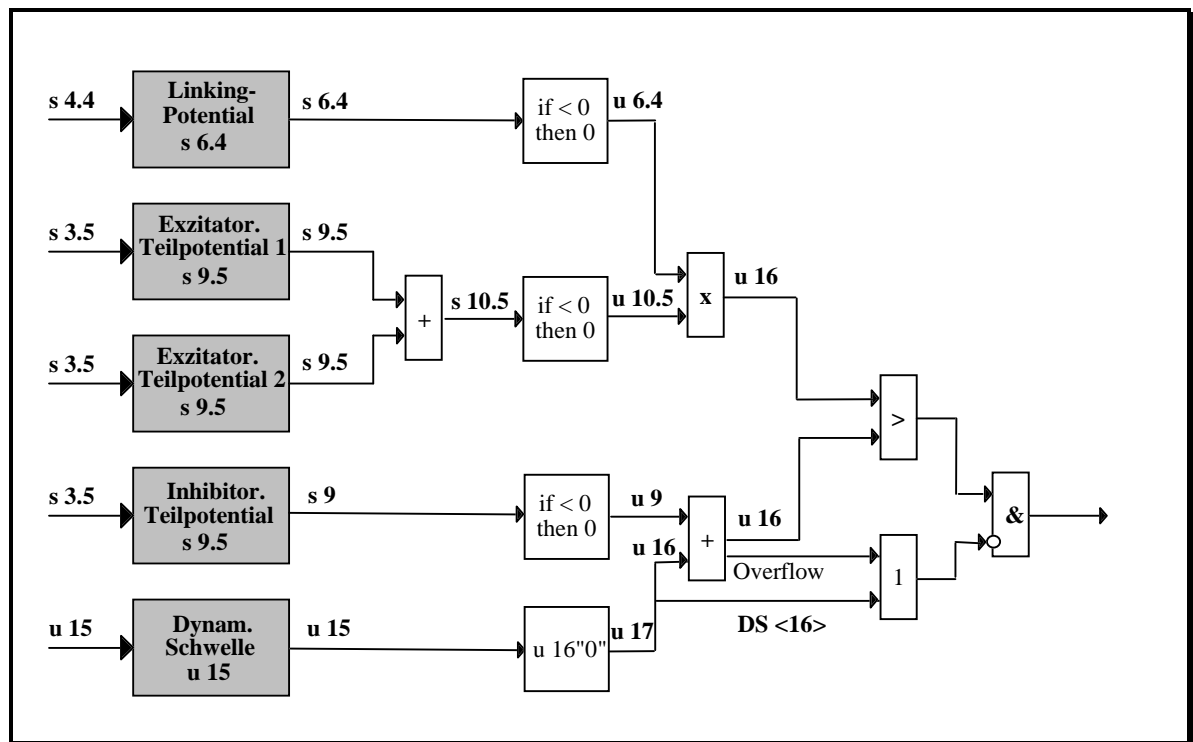


Abb. 2.25: Berechnung des Membranpotentials. Dargestellt sind die Teilpotentiale und die dynamische Schwelle mit ihren Auflösungen und möglichen Gewichten sowie die Verschaltung zur Membranpotentialberechnung und dem Vergleich mit der dynamischen Schwelle. Modifiziert nach Frank et al. (1995).

3 Modelle und Simulationsergebnisse

3.1 Die retinale Vorverarbeitung

3.1.1 Das Vorverarbeitungsmodell

3.1.1.1 Zeitliche Interpolation der Kamerabilder

Die Eingabe des Systems liegt als pixelorientierte Grauwertbildsequenz \mathbf{P}^{Seq} vor. Sie kann direkt von der Kamera oder als eine bereits aufgezeichnete Sequenz aus einem Massenspeicher zugeführt werden. Da, wie bereits in Kapitel 2.3.2 ausgeführt, ein Simulationsschritt Δt einer Millisekunde entspricht, die Kamera aber nur in Zeitabständen von $T^{\text{Kam}} = 40 \text{ ms}$ Bilder liefert, muß zwischen diesen Bildern zeitlich interpoliert werden.

Die Beibehaltung des Grauwertes eines Bildpunktes zwischen zwei aufeinanderfolgenden Bildern würde Unstetigkeiten im zeitlichen Verlauf zur Folge haben, was zu einem künstlich erzeugten, von der Abtastrate der Kamera abhängigen Takt führen würde (Eckhorn 1994). Eine zeitliche Interpolation entsprechend dem Abtasttheorem mit einer $\sin t / t$ -Funktion wäre die ideale Lösung, aber unter Berücksichtigung kurzer Verarbeitungszeiten zu rechenintensiv. Außerdem würde es durch die notwendige Verwendung mehrerer Abtastpunkte zu einem unerwünschten zeitlichen Versatz kommen.

Aus diesen Gründen wird die weniger rechenaufwendige lineare Interpolation benutzt. Hierfür werden lediglich zwei aufeinanderfolgende Abtastpunkte gebraucht. Der resultierende Grauwert $I_{x,y}(t)$ eines Bildpunktes am Ort (x, y) zum Zeitpunkt $t = n \cdot \Delta t$ ergibt sich somit zu

$$I_{x,y}(n \cdot \Delta t) = P_{x,y}(p) + q \cdot \Delta I_{x,y}(p) \quad (\text{Gl. 3.1})$$

mit

$$\Delta I_{x,y}(p) = \frac{P_{x,y}(p+1) - P_{x,y}(p)}{T},$$

$$p = n / T, \quad q = n \bmod T, \quad n \geq 0.$$

Die Variablen p und $q \in \mathbb{N}$ enthalten das Ergebnis der Division mit Rest. $\Delta I_{x,y}(p)$ gibt die Änderung des Grauwertes eines Bildpunktes am Ort (x, y) in einer Millisekunde zwischen zwei aufeinanderfolgenden Bildern $\mathbf{P}(p)$ und $\mathbf{P}(p+1)$ an. T ist der zeitliche Abstand zwischen zwei aufeinanderfolgenden Bildern und muß nicht gleich T^{Kam} sein. Vielmehr erlaubt die Wahl von $T < T^{\text{Kam}}$ eine Verringerung der Rechenzeit. Da der Hardware-Accelerator auf einer 16-Bit-Verarbeitung basiert (vgl. Kap. 2.3.3), wird im folgenden $T = 32 \text{ ms}$ gesetzt. Der resultierende zeitliche Verlauf des Grauwertes eines Bildpunktes $I(t)$ enthält somit keine Unstetigkeiten. Allerdings beinhaltet sein Verlauf noch Stellen, die nicht stetig differenzierbar sind.

3.1.1.2 Hexagonale Unterabtastung der Eingabebilder

Da nur eine begrenzte Anzahl an Modellneuronen zur Verfügung steht und die Rechenzeit möglichst gering sein soll, dürfen auch in der Vorverarbeitung nur relativ wenige Modellneuronen eingesetzt werden. Entsprechend dem biologischen Vorbild, der Fovea centralis, soll nur ein kleiner Teil der Eingabe hochauflösend verarbeitet werden. Das Auflösungsvermögen soll in der Peripherie geringer sein als im Zentrum. Somit muß die Dichte der Rezeptoren im Zentrum hoch, in der Peripherie geringer sein. Ferner soll die Möglichkeit zur leichten Parallelisierung der Vorverarbeitung erhalten bleiben. Aus diesem Grund wurde ein Modell mit mehreren Auflösungsstufen gleicher Netzwerktopologie gewählt. Jede Auflösung enthält die gleiche Anzahl an retinotop angeordneten Verarbeitungseinheiten wie Rezeptoren, Bipolar- und Ganglienzellen, die einen für ihre Auflösung charakteristischen Abtastbereich überdecken. So überdecken die hochauflösendsten Verarbeitungseinheiten den kleinsten Bildbereich. Mit gröber werdender Auflösung wächst der Abstand der Verarbeitungseinheiten und demzufolge auch der zu verarbeitende Bildausschnitt. Die Abbildung 3.1 zeigt schematisch die möglichen Abtastbereiche drei verschiedener Auflösungen.

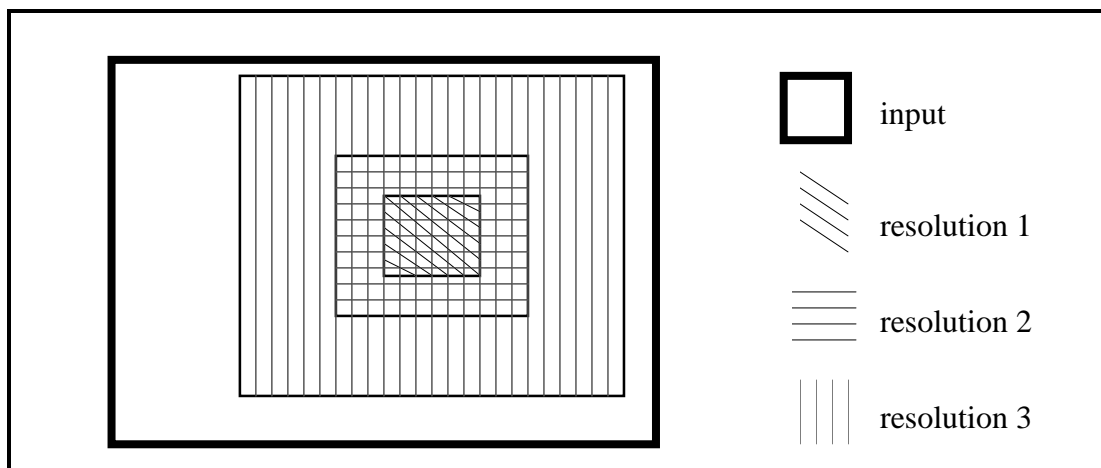


Abb. 3.1: Abtastbereiche der Verarbeitungseinheiten unterschiedlicher Auflösungen. Die Abbildung zeigt die sich überlappenden Abtastbereiche der verschiedenen Auflösungen. Das Eingabebild wird durch die dicken Linien, die Abtastbereiche durch die etwas dünneren Linien begrenzt. Die höchste Auflösung in der Mitte hat den kleinsten Abtastbereich und wird durch die beiden anderen Auflösungen überdeckt.

3.1.1.3 Das Rezeptormodell

Die Empfindlichkeitscharakteristik eines Rezeptors kann nach Rodieck (1965) mit einer zweidimensionalen Gaußverteilung modelliert werden. Zentral einfallende Lichtstrahlen erzeugen ein größeres absolutes Rezeptorpotential als peripher einfallende. Wird die zweidimensionale Gaußverteilung durch eine $(m \times m)$ -Gewichtsmatrix \mathbf{W}^R dargestellt (m ungerade), dann kann das Potential R eines Rezeptors als Faltung der interpolierten Eingabe I mit dieser Gewichtsmatrix dargestellt werden.

$$R_{x,y}(t) = \sum_{i=1}^m \sum_{j=1}^m w_{ij}^R \cdot I_{x+j-(m+1)/2, y+i-(m+1)/2}(t), \quad (\text{Gl. 3.2})$$

mit den Elementen w_{ij} der Matrix \mathbf{W}_R

$$w_{ij}^R = \frac{w'_{k=i, l=j}}{\sum_{k=1}^m \sum_{l=1}^m w'_{kl}} \quad \forall 1 \leq i, j \leq m, \quad (\text{Gl. 3.2b})$$

$$w'_{kl} = \frac{1}{\sqrt{2\pi\sigma_R^2}} \exp\left(-\frac{(k - (m+1)/2)^2 + (l - (m+1)/2)^2}{2\sigma_R^2}\right) \quad \forall 1 \leq k, l \leq m,$$

$$\sum_{i=1}^m \sum_{j=1}^m w_{ij}^R = 1. \quad (\text{Gl. 3.3})$$

Die Matricelemente w_{ij}^R werden durch Gleichung 3.2b so normiert, daß Gleichung 3.3 gilt. Die einzelnen Auflösungen unterscheiden sich durch die Anzahl der Zeilen und Spalten m der Gewichtsmatrix \mathbf{W}^R , der Standardabweichung σ_R der in ihr enthaltenen zweidimensionalen Gaußverteilung und ihrer Dichte. Je größer die Auflösung sein soll, desto größer sind m , σ_R und der Abstand d der Verarbeitungseinheiten zueinander. Hierbei ist zu beachten, daß die Auflösungen nicht durch unterschiedlich konvergente Verschaltungen der Rezeptoren zu Bipolarzellen erzeugt werden. Aus Gründen der Implementierung und Parallelisierbarkeit wird großer Wert auf gleiche Verschaltungen gelegt. Ferner ist das Rezeptorpotential bei Aktivierung positiv und nicht wie bei biologischen Rezeptoren negativ (Hyperpolarisation). Die Anzahl der Zeilen und Spalten m der Matrix \mathbf{W}^R ist als Kompromiß zwischen Rechenzeit – die Anzahl der Multiplikationen bei der Berechnung des Rezeptorpotentials R ist proportional zu m^2 – und den möglichen Fehlern durch die Fensterung der Gaußfunktion mit einer Rechteckfunktion zu sehen.

Die Rezeptoren sind in der dichtmöglichsten Zylinderpackung angeordnet. Daraus resultiert ein quasi hexagonales Raster (Siminoff 1980, Hartmann 1982). Wird das Raster als Matrix \mathbf{Q} dargestellt, dann befinden sich Rezeptoren an den Orten $(x=j, y=i)$, wenn das entsprechende Element q_{ij} der Matrix ungleich null ist.

$$q_{i,j}^d = \begin{cases} 1, & \text{falls } (i_a \leq i \leq i_b) \wedge (j_a \leq j \leq j_b) \wedge \\ & \left((((j - j_a) \bmod d) = 0) \wedge (((i - i_a) \bmod (2 \cdot d)) = 0) \right) \vee \\ & \left((((j - j_a) \bmod d) = d/2) \wedge (((i - i_a) \bmod (2 \cdot d)) = d) \right) \end{cases} \quad (\text{Gl. 3.4})$$

sonst

Der Operator ‘=’ in Gleichung 3.4 ist der logische Vergleichsoperator. Die Parameter i_a , i_b und j_a , j_b geben die erste (a) und letzte (b) Zeile bzw. Spalte des Abtastbereiches der Bipolar- und Ganglienzellen einer Auflösung mit der Abtastweite d an. Um Randeffekte zu vermeiden, sind für die Rezeptoren an allen Seiten des Abtastbereiches zusätzliche Abtastzeilen und Abtastspalten hinzuzufügen. In der Abbildung 3.2 sind die Abtastpunkte für eine Auflösung dargestellt. Jede graue Markierung gibt den Mittelpunkt des rezeptiven Feldes eines Rezeptors an.

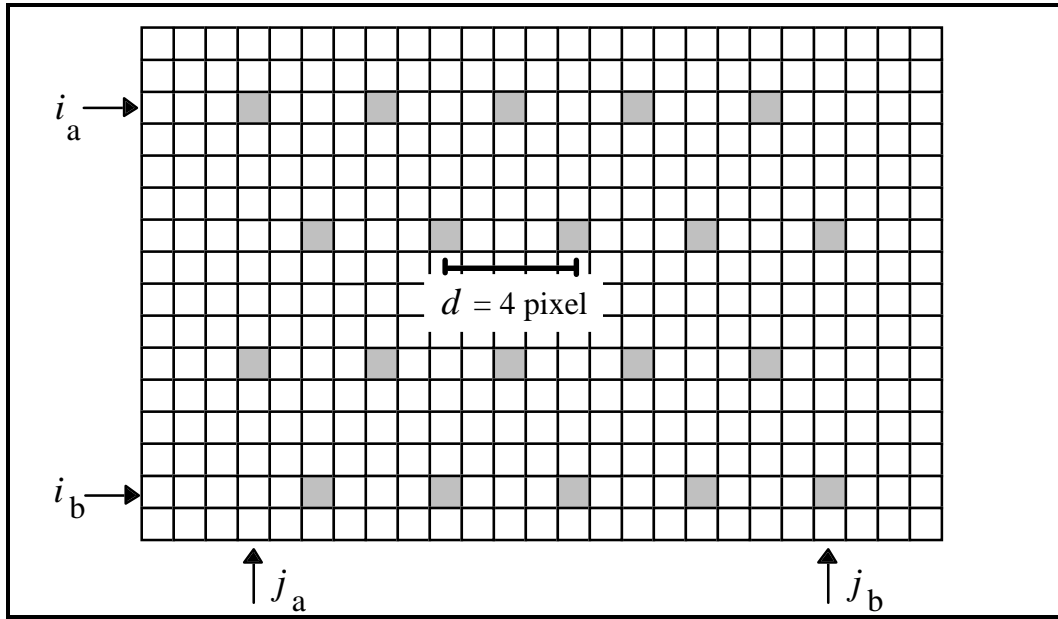


Abb. 3.2: Hexagonales Abtastraaster. Dargestellt ist die hexagonale Abtastung mit einer Abtastweite von $d = 4$ Pixel. Die Kästchen symbolisieren die einzelnen Bildpunkte der Eingabe, die grauen Kästchen die RF-Mittelpunkte der Bipolar- und Ganglienzellen.

Für drei verschiedene Auflösungen werden die Abtastweiten $d = 2, 4$ und 8 Pixel gewählt. Die Standardabweichungen $\sigma_{R,d}$ werden so gewählt, daß jeder Bildpunkt mit der gleichen Gewichtung in die Verarbeitung eingeht. Dadurch überlappen sich die Gewichtsmatrizen zweier benachbarter Rezeptoren ungefähr bei ihrer halben Amplitude. Die Parameter für die drei unterschiedlichen Auflösungen sind in Tabelle 3.1 aufgelistet (Spengler 1996).

Abtastweite d/Pixel	Dimension \mathbf{W}^R	Standardabweichung σ_R
2	5 x 5	1,05
4	11 x 11	2,1
8	23 x 23	4,2

Tab. 3.1: Parameter der Gewichtsmatrizen \mathbf{W}^R . Die Abtastweite d gibt den Abstand zweier benachbarter Verarbeitungseinheiten einer Zeile in Pixel an. Für jede Auflösung wird eine unterschiedlich große Standardabweichung σ_R benutzt, damit jeder Bildpunkt mit der gleichen Gewichtung in die Verarbeitung eingeht. Die Anzahl der Zeilen und Spalten der Gewichtsmatrix \mathbf{W}^R müssen dementsprechend gewählt werden.

3.1.1.4 Das Bipolarzellenmodell

In Kap. 2.2.2 wurde die Verschaltung der Bipolarzellen mit Rezeptoren und Horizontalzellen in der Retina dargestellt. Das rezeptive Feld einer Bipolarzelle besitzt eine antagonistische Zentrum-Umfeld-Struktur. Der Einfluß des Umfeldes wird hierbei durch Horizontalzellen bewerkstelligt. Die Zentrum-Umfeld-Struktur des rezeptiven Feldes der Bipolarzellen eignet sich zur Extraktion lokaler Intensitätskontraste.

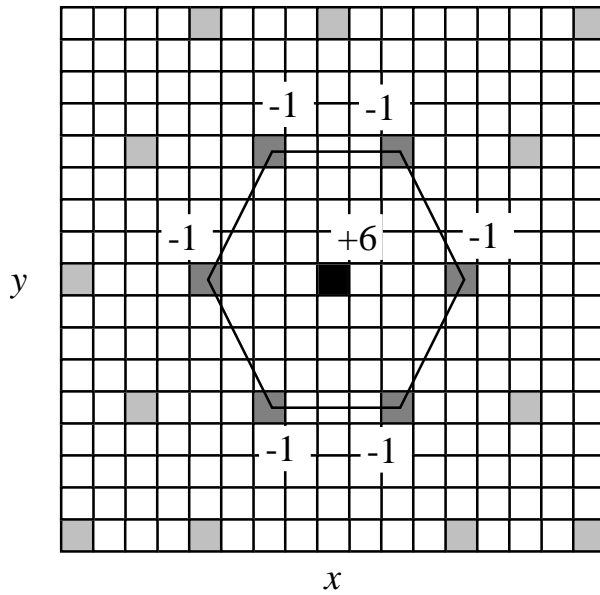


Abb. 3.3: Verschaltung der Rezeptoren zu einer ON-Bipolarzelle. Die Abbildung zeigt einen Ausschnitt des hexagonalen Abtastasters (Abtastweite $d = 4$ Pixel). Das Bipolarzellenpotential $B_{x,y}$ (schwarze Markierung) ergibt sich aus der gewichteten Summation des Rezeptorpotentials $R_{x,y}$ (schwarz) und der Potentiale der benachbarten Rezeptoren des ersten Hexagons (dunkelgrau). Für eine ON-Bipolarzelle wird das Rezeptorpotential $R_{x,y}$ sechsfach positiv, das Potential der sechs benachbarten Rezeptoren einfach negativ gewichtet. Eine OFF-Bipolarzelle ergibt sich durch Vertauschen der Vorzeichen.

Technisch lassen sich Intensitätskontraste durch den sogenannten „Mexikaner-Hut“-Filter, der sich aus der zweifachen Ableitung einer zweidimensionalen Gaußfunktion ergibt, extrahieren. Dieser Filter kann durch Superposition von Gaußverteilungen approximiert werden (Hartmann 1982) und wird nach Young (1987) mit DOOG (englisch: *Difference Of Offset Gaussians*) bezeichnet.

Zur Bildung einer ON-Bipolarzelle mit Zentrum des rezeptiven Feldes am Ort (x, y) werden vom sechsfach positiv gewichteten Potential des Rezeptors (x, y) die Potentiale der sechs nächsten Rezeptoren subtrahiert. Die Potentiale der benachbarten Rezeptoren können hierbei zu einem Horizontalzellenpotential zusammengefaßt werden. Die Abbildung 3.3 zeigt die Gewichtung der Rezeptorpotentiale zur Erzeugung einer ON-Bipolarzelle mit einem rezeptiven Feld, strukturiert in einen antagonistischen Zentrum-Umfeld-Aufbau.

Zusätzlich werden die Rezeptorpotentiale noch mit einem zeitlichen Tiefpaß gefiltert. Dadurch werden nicht stetig differenzierbare Stellen im zeitlichen Verlauf der Rezeptorpotentiale geglättet. Die Tiefpaßfilterung wird, wie in Kapitel 3.2 beschrieben, als Faltung der Rezeptorpotentiale mit der Stoßantwort $h^B(t)$ eines Tiefpasses implementiert. Das Potential $B_{x,y}^{\text{ON}}$ einer ON-Bipolarzelle mit rezeptiven Feld Zentrum am Ort (x, y) ergibt sich dann aus Gleichung 3.4.

$$B_{x,y}^{\text{ON}}(t) = \left(6 \cdot R_{x,y}(t) - H_{x,y}(t) \right) * h^B(t) \quad (\text{Gl. 3.4})$$

Mit $H(t)$ wird das Potential der Horizontalzelle gekennzeichnet, das sich aus der Summe der sechs am nächsten benachbarten Rezeptorpotentiale, die Potentiale der Rezeptoren des ersten Hexagons, errechnet:

$$H_{x,y}(t) = \sum_{k=-1}^1 \sum_{\substack{l=-1 \\ l \neq 0}}^1 R_{x+l \cdot (d-|k| \cdot d/2), y+k \cdot d} \cdot \quad (\text{Gl. 3.5})$$

Das Potential $B_{x,y}^{\text{OFF}}$ einer OFF-Bipolarzelle ergibt sich aus der Vorzeichenumkehr des ON-Bipolarzellenpotentials $B_{x,y}^{\text{ON}}$.

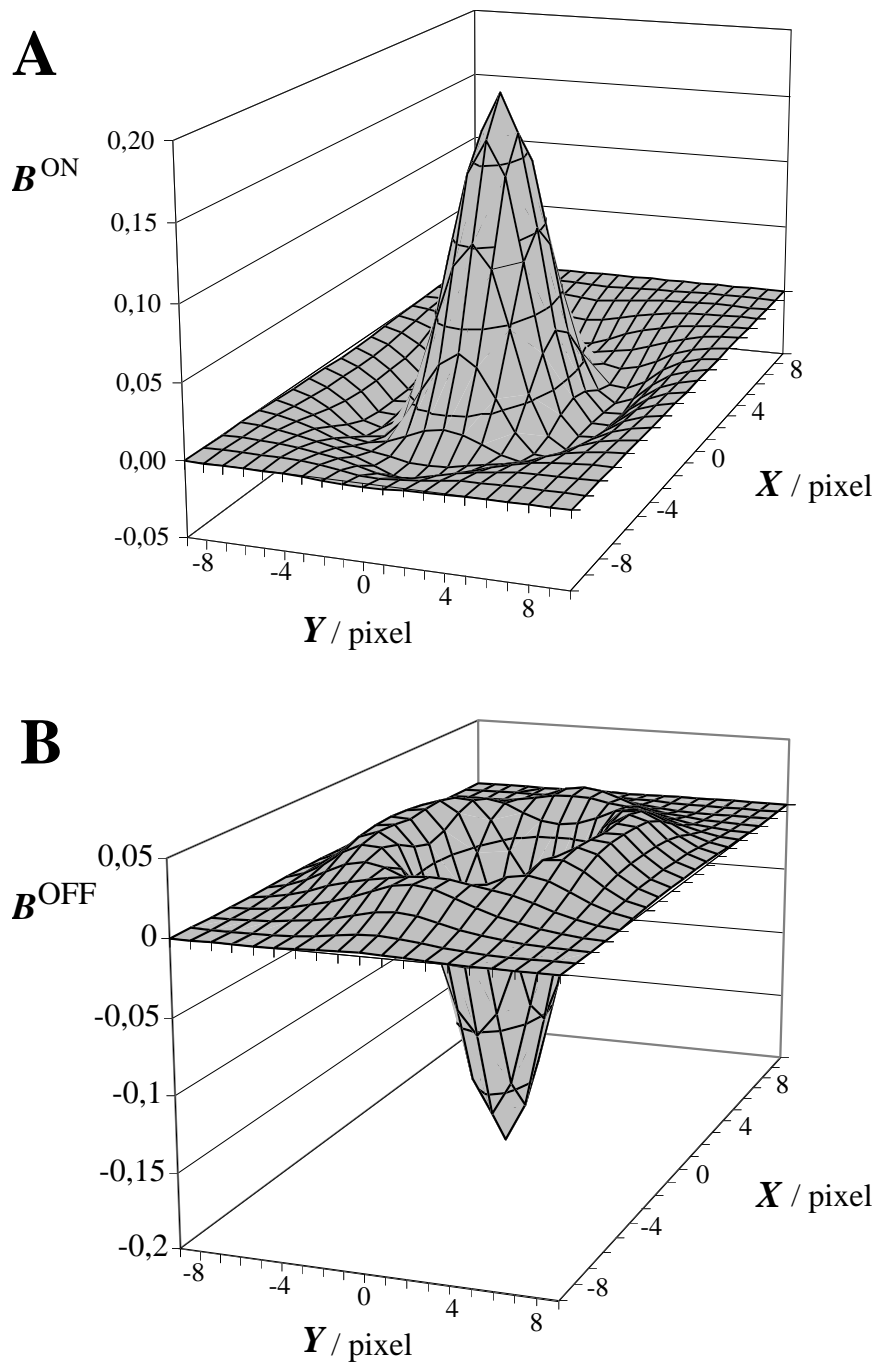


Abb. 3.4: Sensitivitätsprofil der Bipolarzellen. Für die Auflösung mit der Abtastweite $d = 4$ Pixel sind in **A** und **B** die Potentiale B der ON- bzw. OFF-Bipolarzellen, die sich bei Stimulation mit einem Lichtfleck der Größe eines Pixels und des interpolierten Grauwertes $I = 1$ ohne Berücksichtigung der Tiefpaßfilterung ergeben würden, dargestellt.

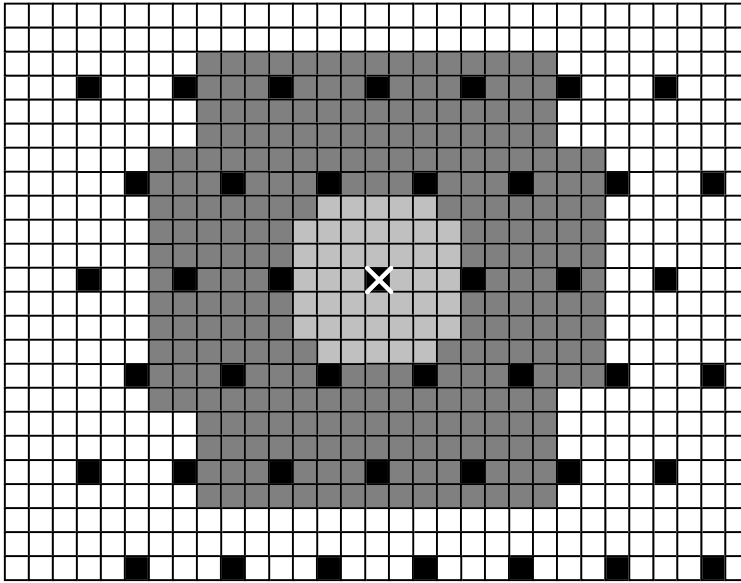


Abb. 3.5: Rezeptives Feld einer Bipolarzelle im Abtasteraster. Der grau markierte Bereich gibt den Bereich des rezeptiven Feldes einer Bipolarzelle (weißes Kreuz) an. Hellgrau ist das Zentrum, dunkelgrau das Umfeld dargestellt. An jedem schwarz markierten Ort ist eine Bipolarzelle mit einem solchen rezeptiven Feld vorhanden. Eine vollständige, homogene Ausleuchtung des rezeptiven Feldes führt zu einem verschwindenden Potential.

$$B_{x,y}^{\text{OFF}}(t) = (-1) \cdot B_{x,y}^{\text{ON}}(t) \quad (\text{Gl. 3.6})$$

Die in Abbildung 3.5 dargestellten rezeptiven Felder der Bipolarzellen könnten auch direkt durch eine Faltung des interpolierten Grauwertbildes \mathbf{I} mit einer Gewichtsmatrix \mathbf{W}^B , deren Elemente den entsprechenden Potentialwerten aus Abbildung 3.4 entsprächen, erzeugt werden. Hierbei würde allerdings die Möglichkeit entfallen, unterschiedliche zeitliche Eigenschaften bei der Signalübertragung von Rezeptor zu Bipolarzelle oder Horizontalzelle zu Bipolarzelle zu wählen. Weiterhin wäre hierfür eine größere Anzahl an Multiplikationen erforderlich.

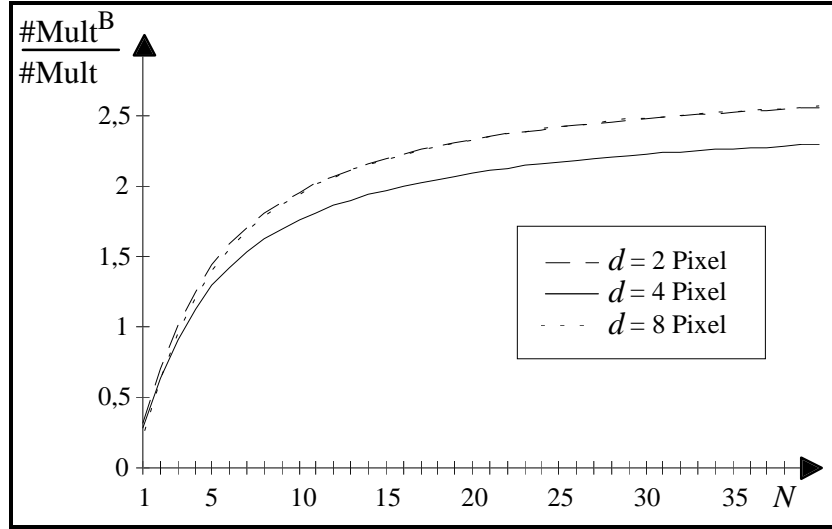
Bei Verwendung der Gewichtsmatrix \mathbf{W}^B läßt sich die Anzahl der Multiplikationen $\#Mult^B$ durch Gleichung 3.7 ausdrücken. Die Gleichung 3.8 gibt im Vergleich dazu die Anzahl der Multiplikationen $\#Mult$ bei der in dieser Arbeit verwendeten Verschaltung an.

$$\#Mult^B = (2 \cdot d + m)^2 \cdot N \cdot M \quad (\text{Gl. 3.7})$$

$$\#Mult = (N + 2) \cdot (M + 2) \cdot (m^2 + 7) \quad (\text{Gl. 3.8})$$

In den beiden Gleichungen bezeichnen N und M die Anzahl der Abtastpunkte in horizontaler bzw. vertikaler Richtung, m steht für die Anzahl der Zeilen und Spalten der Gewichtsmatrix \mathbf{W}^B bei der Auflösung mit Abtastweite d . In Abb. 3.6 sind die Quotienten aus $\#Mult^B$ und $\#Mult$ für die drei Auflösungen mit den Parametern aus Tab. 3.1 unter der Vereinfachung $M=N$ aufgetragen. Nur bei sehr kleinen Netzen ($N < 5$) ist die direkte Berechnung des Bipolarzellenpotentials mit der Gewichtsmatrix \mathbf{W}^B von Vorteil. Bei einem Netz mit der Größe $N=20$ werden zum Beispiel mehr als doppelt so viele Multiplikationen benötigt. Die Ersparnis liegt an der mehrfachen Verwendung der Rezeptorpotentiale.

Abb. 3.6: Anzahl der Multiplikationen bei der Berechnung des Bipolarzellenpotentials. In der Abbildung ist das Verhältnis der notwendigen Multiplikationen bei direkter Berechnung des Bipolarzellenpotentials und bei indirekter Berechnung mit Hilfe von Rezeptoren in Abhängigkeit von der Anzahl an Abtastpunkten in horizontaler und vertikaler Richtung aufgetragen. Nur bei sehr kleinen Netzen ist die direkte Berechnung des Bipolarzellenpotentials von Vorteil.



3.1.1.5 Das Ganglienzellmodell

Das zeitliche Antwortverhalten einer Ganglienzelle setzt sich aus einem phasischen (zeitlicher Bandpaß) und einem tonischen Teil (zeitlicher Tiefpaß) zusammen. Bei Ganglienzellen, die hauptsächlich auf Transienten reagieren, ist der phasische Anteil deutlich ausgeprägter als der tonische. Da das zu simulierende Gesamtsystem zeitliche Änderungen des Eingabebildes verarbeiten soll, werden hier nur Transientenzellen, vergleichbar mit den Y-Ganglienzellen der Katze oder den M-Ganglienzellen des Affen, modelliert. Da Ganglienzellen Aktionspotentiale erzeugen, ist das Marburger Modellneuron Grundlage des Ganglienzellmodells.

Die Transientendetektion läßt sich durch einen exzitatorischen und einen inhibitorischen Dendriten mit unterschiedlichen zeitlichen Übertragungseigenschaften bewerkstelligen. Der Leckintegrator des exzitatorischen Dendriten zeichnet sich im Gegensatz zu dem des inhibitorischen Dendriten durch eine kleine Zeitkonstante aus. Die Potentiale der Ganglienzellen ergeben sich zu

$$G_{x,y}(t) = G_{x,y}^{exz}(t) + G_{x,y}^{inh}(t) \quad (\text{Gl. 3.9})$$

mit

$$\begin{aligned} G_{x,y}^{exz}(t) &= y_{x,y}^B(t) * h^{G,exz}(t) \\ G_{x,y}^{inh}(t) &= y_{x,y}^B(t) * h^{G,inh}(t) \\ y_{x,y}^B(t) &= \begin{cases} B_{x,y}(t), & \text{für } B_{x,y}(t) > 0 \\ 0, & \text{sonst} \end{cases} \end{aligned} \quad (\text{Gl. 3.10})$$

Die Anwendung der Gleichung 3.10 ist notwendig, damit eine Hyperpolarisation der Bipolarzelle nicht zu einer Depolarisation der Ganglienzelle über den inhibitorischen Dendriten führt. Ein Abfall des ON-Bipolarzellenpotentials geht gemäß Gleichung 3.6 immer mit einem Anstieg des OFF-Bipolarzellenpotentials einher. Ohne Gleichung 3.10 wäre nicht

mehr eindeutig feststellbar, ob die Depolarisation der Ganglienzelle von einem Anstieg oder Abfall im Bipolarzellenpotentialverlauf herrührt.

Die Leckintegratoren werden durch die Parameter $V^{G,exz}$, $\tau^{G,exz}$, $V^{G,inh}$ und $\tau^{G,inh}$ charakterisiert (vgl. Kap. 2.3.2). Die Parameter werden so gewählt, daß nur Anstiege im Potentialverlauf der Bipolarzelle zu einer Depolarisation der Ganglienzelle führen. Deswegen muß das exzitatorische postsynaptische Potential im Vergleich zum inhibitorischen schnell ansteigen. Dazu müssen die Zeitkonstanten $\tau^{G,exz}$ und $\tau^{G,inh}$ die Bedingung $\tau^{G,exz} < \tau^{G,inh}$ erfüllen. Die Verstärkungsfaktoren $V^{G,exz}$ und $V^{G,inh}$ sind nicht frei wählbar, sondern müssen zu einem minimalen tonischen Gesamtpotential führen. Für den konstruierten Fall eines Potentialverlaufs einer Bipolarzelle in Form einer Sprungfunktion $i(t)$ mit

$$i(t) = \begin{cases} 1, & t \geq 0 \\ 0, & t < 0 \end{cases} \quad (\text{Gl. 3.11})$$

läßt sich eine Randbedingung für die Wahl der Verstärkungsfaktoren angeben. Für das Potential der Ganglienzelle gelten die Gleichungen 3.12 und 3.13.

$$G(t+1) = G(t) + \Delta G(t) \quad (\text{Gl. 3.12})$$

$$G(t+1) = G(t) \cdot \exp(-\Delta t / \tau^G) + i(t) \cdot V^G \quad (\text{Gl. 3.13})$$

In Gleichung 3.12 gibt $\Delta G(t)$ die Änderung des Ganglienzellenpotentials im diskreten Zeitschritt t an. Hat $G(t)$ das Sättigungspotential erreicht, dann ist $\Delta G(t)$ gleich null, und es ergibt sich durch Einsetzen von Gleichung 3.12 in Gleichung 3.13 das Potential der Ganglienzelle in der Sättigung zu

$$G^{\text{Sätt}} = \frac{V^G}{(1 - \exp(-\Delta t / \tau^G))}. \quad (\text{Gl. 3.14})$$

Soll das resultierende Gesamtpotential minimal werden, muß

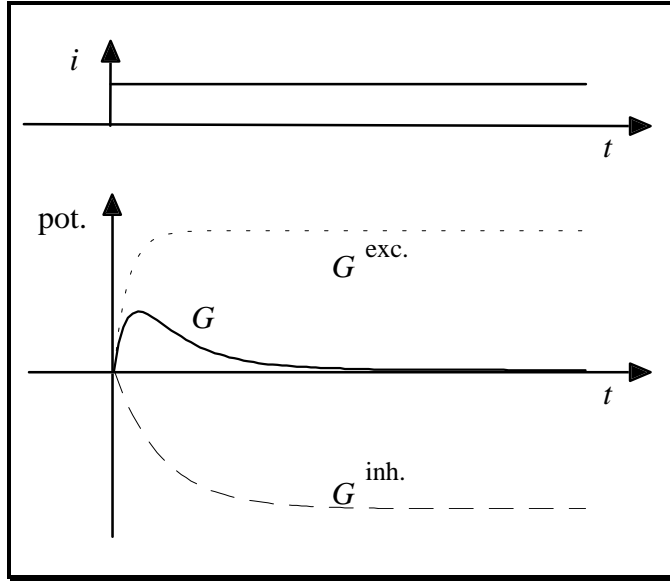
$$G^{\text{Sätt},exz} + G^{\text{Sätt},inh} = 0 \quad (\text{Gl. 3.15})$$

sein. Damit ergibt sich die gesuchte Randbedingung für ein minimales tonisches Gesamtpotential der Ganglienzelle zu

$$\frac{V^{G,exz}}{(1 - \exp(-\Delta t / \tau^{G,exz}))} = \frac{-V^{G,inh}}{(1 - \exp(-\Delta t / \tau^{G,inh}))}. \quad (\text{Gl. 3.16})$$

Bei Vorgabe von $V^{G,exz}$, $\tau^{G,exz}$ und $\tau^{G,inh}$ ist $V^{G,inh}$ gemäß Gleichung 3.16 zu wählen, damit das resultierende Gesamtpotential der Ganglienzelle bei einem konstanten Eingangspotential minimal wird. Durch einen etwas größer gewählten Verstärkungsfaktor V^{exz} wird das Gesamtpotential nicht null, sondern bleibt positiv. Dadurch kann der tonische Teil des Potentials modelliert werden. Eine solche Modellierung würde einer X-Ganglienzelle entsprechen. Der Verlauf des Gesamtpotentials G einer Ganglienzelle mit geringem tonischen Anteil ist schematisch in Abb. 3.7 für die Eingabe nach Gleichung 3.11 dargestellt.

Abb. 3.7: Schematischer Verlauf des Membranpotentials an Dendriten und Spike-Encoder Eingang einer Ganglienzelle. Ein Stimulus, der zur Zeit $t=0$ von $i(t<0)=0$ auf $i(t\geq 0)=1$ gesetzt wird, führt zu einem absoluten Potentialanstieg in den beiden Dendriten. Aufgrund der Parameter steigt das exzitatorische Potential G^{exc} schneller als das inhibitorische G^{inh} . Dadurch wird das Gesamtpotential G kurzzeitig positiv, die Ganglienzelle aktiv und der Anstieg in der Eingabe somit detektiert. Da die Absolutwerte der beiden Sättigungspotentiale gleich sind, fällt das Gesamtpotential auf null ab.



Damit das Potential G nicht zu stark von dem detektierten Intensitätskontrast abhängt, wird es zusätzlich noch mit einer Sigmoidfunktion gestaucht. Das Potential der Ganglienzelle G' , das zum Schwellwertvergleich herangezogen wird, berechnet sich dann zu

$$G'(t) = G^{\max} \cdot \left(\frac{1}{(1 + \exp(-\alpha \cdot G(t)))} - \frac{1}{2} \right). \quad (\text{Gl. 3.17})$$

Der Parameter G^{\max} gibt den maximalen Grauwert an. Die Steigung der Sigmoidkurve wird durch den Parameter α bestimmt.

Die Ausgabe $y^G(t)$ einer Ganglienzelle zum Zeitpunkt t berechnet sich dann nach Gleichung 3.18 zu

$$y^G(t) = \begin{cases} 1, & \text{falls } G'(t) \geq (\Theta^G(t) + \Theta_0^G) \\ 0, & \text{sonst} \end{cases} \quad (\text{Gl. 3.18})$$

mit

$$\Theta^G(t) = y^G(t) * h^{G,\Theta}(t).$$

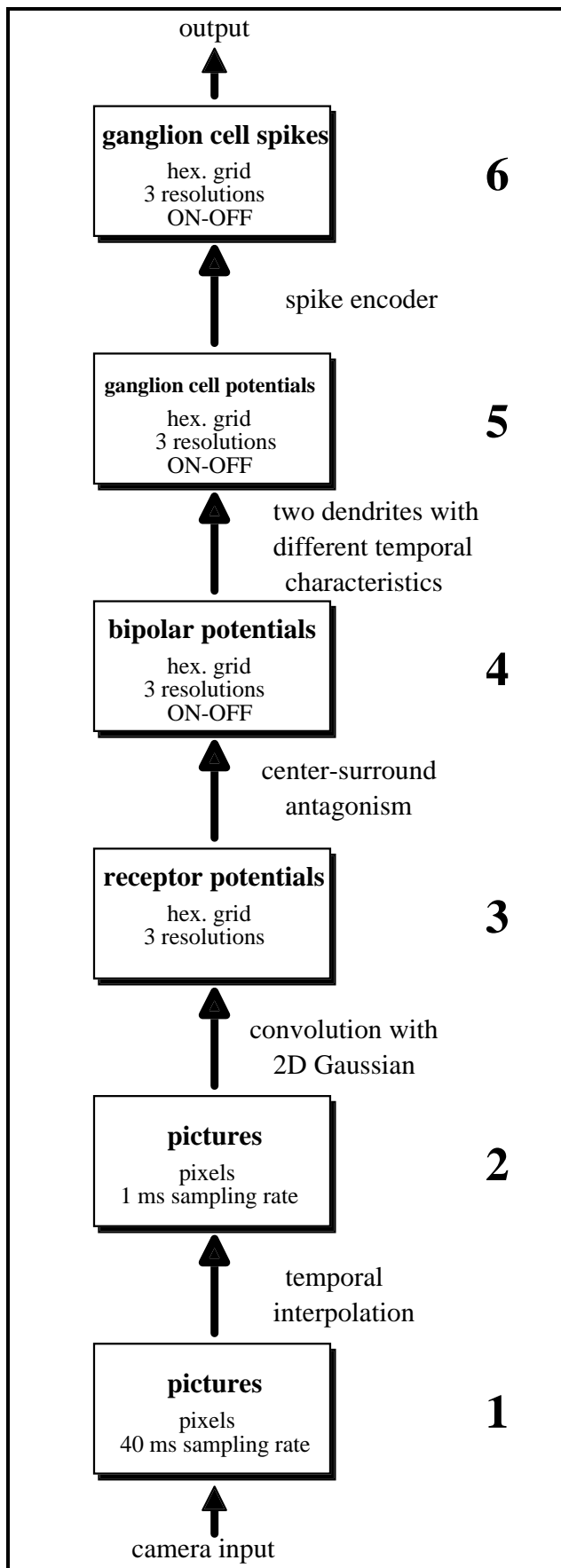


Abb. 3.8: Übersicht Vorverarbeitungsmodell.

(1) Kamerabilder liegen im 40 ms-Takt aufgenommen als Eingabe vor. (2) Nach der zeitlichen Interpolation liegt in jeder Millisekunde ein Bild als Eingabe vor. (3) Diese Bilder werden in 3 Auflösungen hexagonal unterabgetastet und mit Gaußmasken gemäß Tabelle 3.1 gefaltet. (4) Die so entstandenen Rezeptorpotentiale werden über eine DOOG-Architektur zu ON- und OFF-Bipolarzellen verschaltet. (5) Die Bipolarzellpotentiale werden jeweils über einen exzitatorischen und einen inhibitorischen Dendriten mit unterschiedlichen zeitlichen Eigenschaften zu Ganglienzellpotentialen verrechnet. (6) Der Spike-Encoder des Marburger Modellneurons setzt die Ganglienzellpotentiale in Aktionspotentiale um, die ihrerseits die Eingabe des Acceleratorsystems darstellen.

3.1.2 Simulationsergebnisse Vorverarbeitung

3.1.2.1 Simulationsergebnisse mit künstlichen Szenen

In diesem Kapitel wird das Übertragungsverhalten der Vorverarbeitungseinheiten auf sich bewegende künstliche Stimuli untersucht. Als Netzwerkeingabe dienen eine gerade Intensitätskante, ein Balkengitterreiz und eine räumlich sinusförmige (eindimensionale) Intensitätsverteilung. Die Simulationen mit der Kante und der sinusförmigen Intensitätsverteilung zeigen die Aktivierung der Vorverarbeitungseinheiten in Abhängigkeit von Geschwindigkeit und Grauwert bzw. „Kantenschärfe“. Anhand des Gitterreizes wird überprüft, ob die retinalen Ganglienzellen in der Lage sind, auch schnell aufeinanderfolgende Intensitätskontrastkanten zu verarbeiten.

Für die in Kapitel 3.1.1 beschriebene Verschaltungsstruktur der Bildvorverarbeitung werden die Parameter aus Tabelle 3.2 verwendet. Die Parameter sind für alle Auflösungen sowie ON- und OFF-Pfad gleich. Alle Simulationen dieser Arbeit werden mit diesen Vorverarbeitungsparametern durchgeführt.

In der Bildvorverarbeitung stellen die Rezeptoren die ersten Verarbeitungseinheiten dar. Sie führen eine räumliche Tiefpaßfilterung der Eingangsbilder durch und geben ihre daraus resultierenden Potentiale zeitlich tiefpaßgefiltert auf die Bipolarzellen weiter. Zur Untersuchung dieser Potentialverläufe wird ein Stimulus in Form einer geraden, 90° orientierten Kante benutzt. Die von der Kante getrennten Bereiche haben die Grauwerte $g_1=150$ und $g_2=50$ (Bildgröße 128 Pixel x 128 Pixel). Die Kante wird mit unterschiedlichen Geschwindigkeiten v in Richtung 0° bewegt. Die nachfolgenden Simulationsergebnisse verdeutlichen ebenfalls die Abhängigkeit der Potentiale von der genauen Lage der Kante im rezeptiven Feld. Wird die Kante nur um $\Delta = 1$ Pixel verschoben, so ändern sich bereits die Potentiale der Verarbeitungseinheiten. Dadurch wird es notwendig, die unterschiedlichen Positionen der Kante in die Untersuchungen miteinzubeziehen. Die Anzahl der Pixel, die von der Kante von einem zum nächsten Bild übersprungen werden, entspricht hierbei auch der Anzahl an möglichen Positionsverschiebungen, die zu unterschiedlichen Potentialen führen.

Leckintegrator Bipolarzelle τ^B	10 Bin
Leckintegrator Ganglienzelle, exz. Dendrit $V^{G,exz}$	1,0
Leckintegrator Ganglienzelle, exz. Dendrit $\tau^{G,exz}$	3,0 Bin
Leckintegrator Ganglienzelle, inh. Dendrit $V^{G,inh}$	-0,55
Leckintegrator Ganglienzelle, inh. Dendrit $\tau^{G,inh}$	6,5 Bin
Schwellenverstärkung Ganglienzelle V^θ	35
Schwellenzeitkonstante Ganglienzelle τ^θ	10 Bin
Schwellenoffset Ganglienzelle θ_0	20

Tab. 3.2: Netzwerkparameter Bildvorverarbeitung

Die Abbildung 3.10 zeigt den zeitlichen Verlauf der tiefpaßgefilterten Potentiale eines Rezeptors der mittleren Auflösung ($d = 4$ Pixel) aus der Mitte des Abtastrasters ($N=20, M=10$) für fünf verschiedene Geschwindigkeiten v mit $1 \text{ Pixel/Bild} \leq v \leq 5 \text{ Pixel/Bild}$. Der zeitliche Verlauf unterteilt sich nach dem Anschalteffekt ($t < 50 \text{ Bin}$) in zwei Sättigungsbereiche und einen Bereich, in dem das Potential ansteigt. Der Anstieg ist um so steiler, je höher die Geschwindigkeit ist, da die Verweildauer der Kante im rezeptiven Feld für höher werdende Geschwindigkeiten kleiner wird. Für die Grauwerte g_1 und g_2 ergeben sich die Sättigungspotentiale nach Gleichung 3.14. In diesem Beispiel betragen die Sättigungspotentiale $R^{\text{Sätt},g_1} \approx 790$ und $R^{\text{Sätt},g_2} \approx 260$. Die Steigung des Anstiegs ist somit auch von der Differenz der Grauwerte g_1 und g_2 abhängig.

Die Abbildung 3.11 zeigt den tiefpaßgefilterten Potentialverlauf des gleichen Rezeptors bei unterschiedlichen Positionsverschiebungen Δ , wobei Δ bei einer Kantengeschwindigkeit von $v = 4 \text{ Pixel/Bild}$ die Werte von 0 Pixel bis 3 Pixel annimmt. Die Kurven ähneln sich sehr und unterscheiden sich nur durch einen zeitlich äquidistanten Versatz im Anstieg, der durch die unterschiedliche Position entsteht. Diese Ähnlichkeit im Potentialverlauf tritt nur bei den Rezeptoren auf und trifft für die weiteren Verarbeitungseinheiten nicht mehr zu.

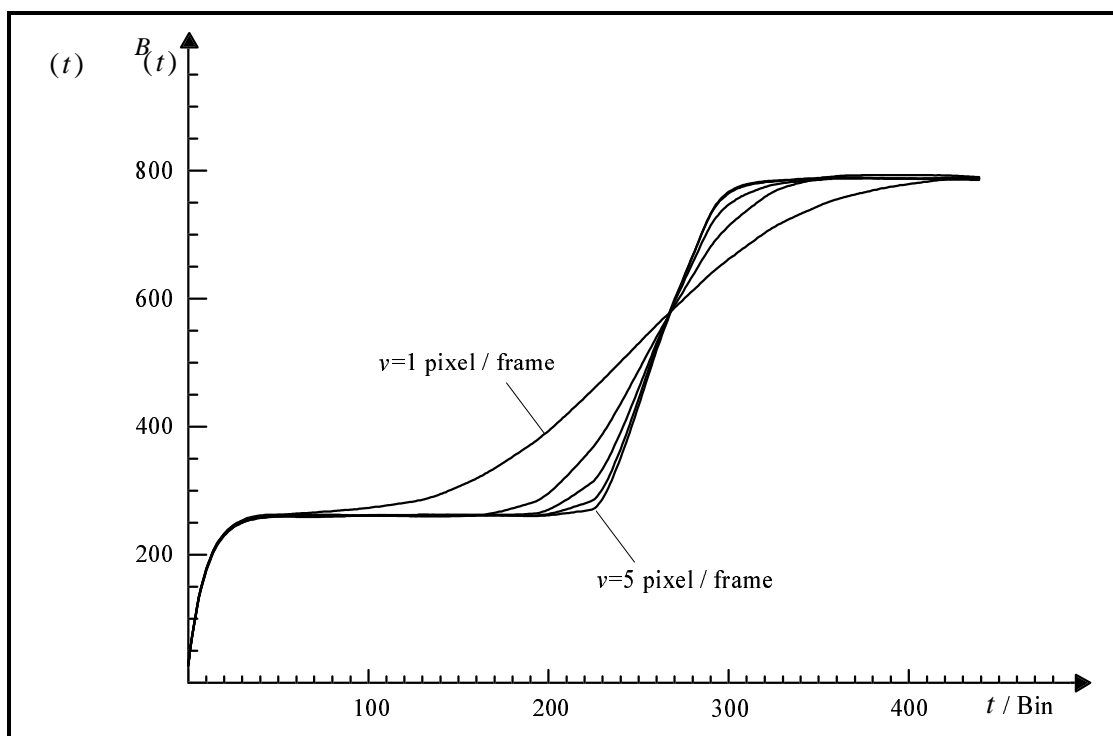


Abb. 3.10: Zeitlicher Verlauf der tiefpaßgefilterten Rezeptorpotentiale für verschiedene Geschwindigkeiten. Für einen in der Mitte des Abtastrasters gelegenen Rezeptor sind die tiefpaßgefilterten Rezeptorpotentiale $R(t) * h^B(t)$ für verschiedene Kantengeschwindigkeiten $v = 1 \text{ Pixel/Bild}$ bis $v = 5 \text{ Pixel/Bild}$ mit $g_1 = 150$, $g_2 = 50$ und $\Delta = 0 \text{ Pixel}$ aufgetragen (Parameter gemäß Tabelle 3.2). Das Rezeptorpotential befindet sich zunächst in einem für den Grauwert g_2 typischen Sättigungswert und steigt dann in den Sättigungswert für den Grauwert g_1 . Der Anstieg ist umso steiler, je größer die Geschwindigkeit ist.

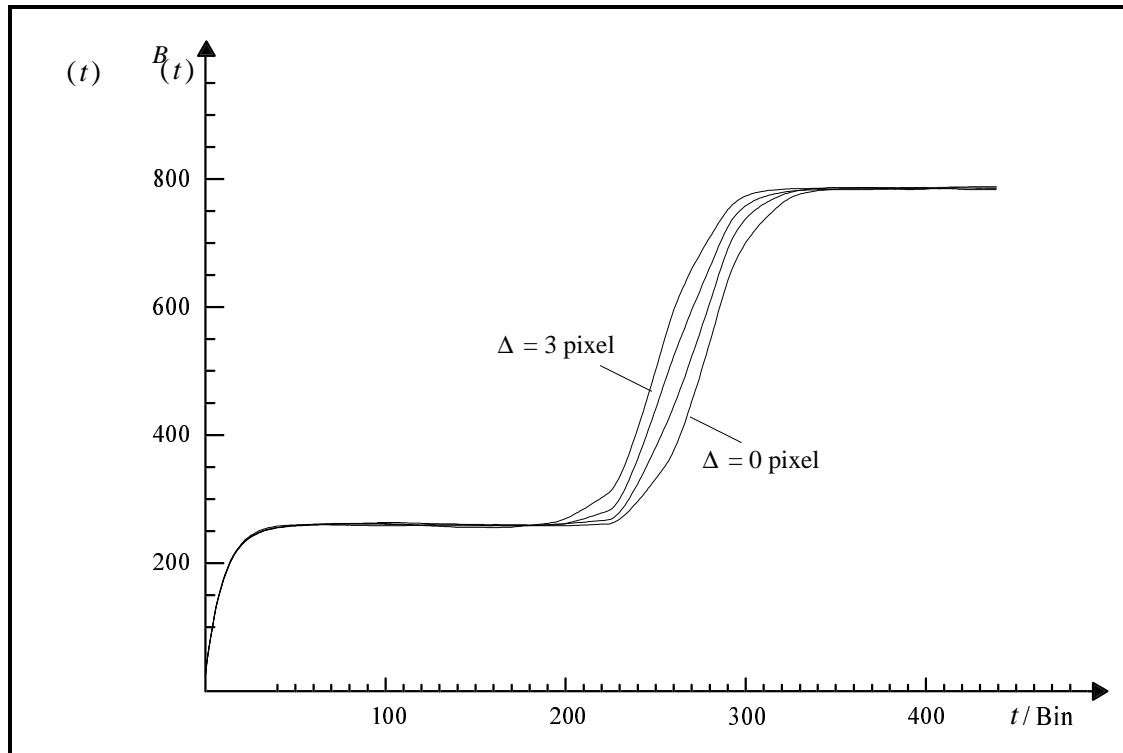


Abb. 3.11: Zeitlicher Verlauf der tiefpaßgefilterten Rezeptorpotentiale für verschiedene Positionsverschiebungen. Eine Kante ($g_1=150$, $g_2=50$) bewegt sich mit $v = 4$ Pixel/Bild über das rezeptive Feld eines Rezeptors. Dargestellt ist das tiefpaßgefilterte Rezeptorpotential $R(t)*h^B(t)$ für vier verschiedene Positionsverschiebungen Δ (Parameter gemäß Tabelle 3.2). Die Kurven haben einen äquidistanten Versatz im Anstiegsbereich.

Die Bipolarzellen haben aufgrund ihrer DOOG-Verschaltungsstruktur räumliche Bandpaßeigenschaften, wodurch sie auf lokale Intensitätskontraste im Eingabebild empfindlich sind. Eine homogene Ausleuchtung des rezeptiven Feldes führt zu einem Bipolarzellenpotential von null, was aus der gleichen Gewichtung von Zentrum und Umfeld in der Verschaltungsstruktur resultiert. Befindet sich, wie in Abb. 3.12 A schematisch gezeigt, eine Kante im rezeptiven Feld einer ON-Bipolarzelle, so kann das Potential positiv oder negativ werden, wenn der größere Teil des RF-Zentrums von der helleren bzw. dunkleren Seite der Kante stimuliert wird. Für eine OFF-Bipolarzelle gilt dies aufgrund der Vorzeichenumkehr nach Gleichung 3.6 umgekehrt. Die Abb. 3.12 B zeigt die Sättigungspotentiale einer OFF-Bipolarzelle ($d = 4$ Pixel) für eine statische, vertikal orientierte Kante ($g_1=100$, $g_2=0$) in Abhängigkeit zur relativen Lage x der Kante zum Mittelpunkt des rezeptiven Feldes.

Liegt der Mittelpunkt des rezeptiven Feldes einer OFF-Bipolarzelle genau auf der hellen Seite der Kante, so wird der größere Teil des Zentrums mit dem Grauwert g_1 stimuliert, was zu einem negativen Potential führt. Rückt die Kante nach links, d.h. die relative Verschiebung x wird negativ, dann ist das Potential solange positiv, bis die Kante das rezeptive Feld verläßt. Eine Verschiebung der Kante nach rechts führt hingegen zu einem negativen Potential.

Das Potentialmaximum einer OFF-Bipolarzelle liegt bei einer nach Tabelle 3.1 gewählten rezeptiven Feldgröße für die mittlere Auflösung ($d = 4$ Pixel) und einer hell/dunkel-Kante ($g_1 > g_2$) bei $x = -3$ Pixel. Das Potential einer ON-Bipolarzelle wird bei $x = 2$ Pixel maximal. Die

RF-Mittelpunkte aktiver OFF-Bipolarzellen liegen somit auf der dunklen Seite einer Kante, die der aktiven ON-Bipolarzellen auf der hellen.

Die Abbildung 3.13 zeigt die zeitlichen Potentialverläufe einer ON-Bipolarzelle für eine sich mit einer Geschwindigkeit v über das receptive Feld bewegende Kante. Die Grauwerte der Kante sind $g_1=150$ und $g_2=50$. Die Kantengeschwindigkeit v liegt zwischen $v = 1$ Pixel/Bild und $v = 6$ Pixel/Bild. Jeder Pixel wird dabei unabhängig von seinen Nachbarn zeitlich verrauscht (normalverteilt mit $\sigma = 1$, Amplitude ± 10).

Alle Potentiale werden zunächst negativ, dann positiv. Dies ist in der Zentrum-Umfeld-Struktur des rezeptiven Feldes begründet. Das receptive Feld wird zunächst nur mit dem Grauwert g_2 stimuliert. Dann läuft von links die hellere Seite der Kante mit dem Grauwert g_1 in das receptive Feld hinein. Da die Kante sich zunächst nur auf dem Umfeld befindet, wird das Potential negativ. Erreicht die Kante das Zentrum, steigt aufgrund der dort befindlichen positiven Gewichte das Potential an. Verläßt die Kante das Zentrum, fällt das Potential wieder ab.

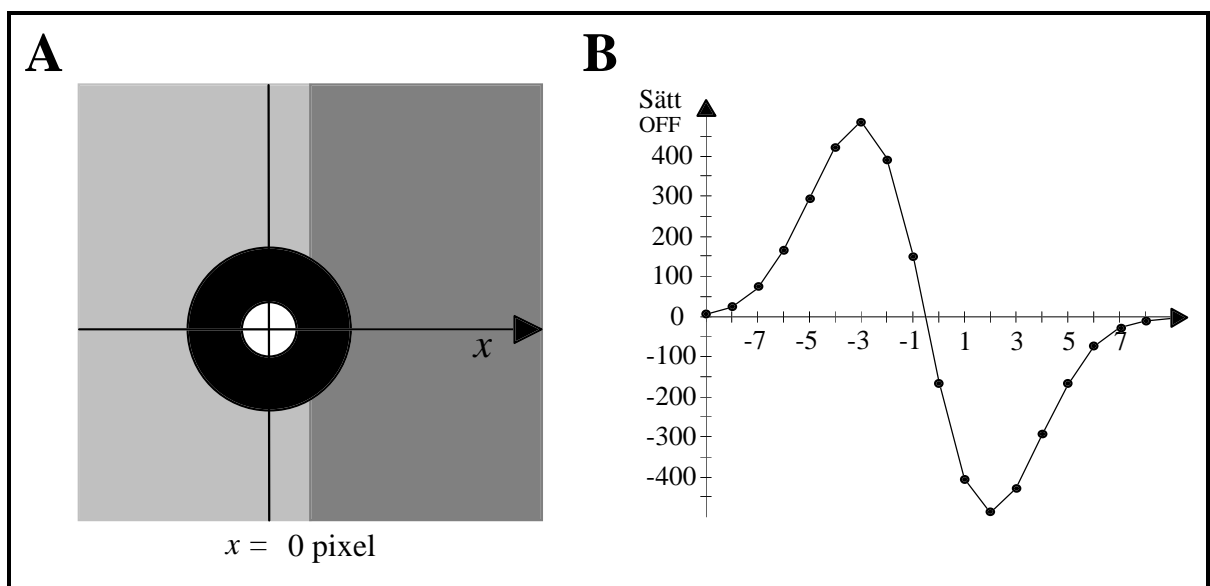


Abb. 3.12: Sättigungspotentiale einer Bipolarzelle bei Stimulation mit einer statischen Kante. A: Schematische Darstellung des rezeptiven Feldes einer ON-Bipolarzelle auf einer Kante. **B:** Sättigungspotentiale einer OFF-Bipolarzelle in Abhängigkeit zur relativen Lage von RF-Mittelpunkt zur Kante. Für $x = 0$ Pixel liegt der Mittelpunkt des rezeptiven Feldes auf der linken Seite der Kante. Nur die schwarzen Kreise geben Sättigungspotentiale für eine bestimmte relative Verschiebung x an. Die durchgezogene Linie dient lediglich der besseren Darstellung. Nähere Erläuterungen im Text.

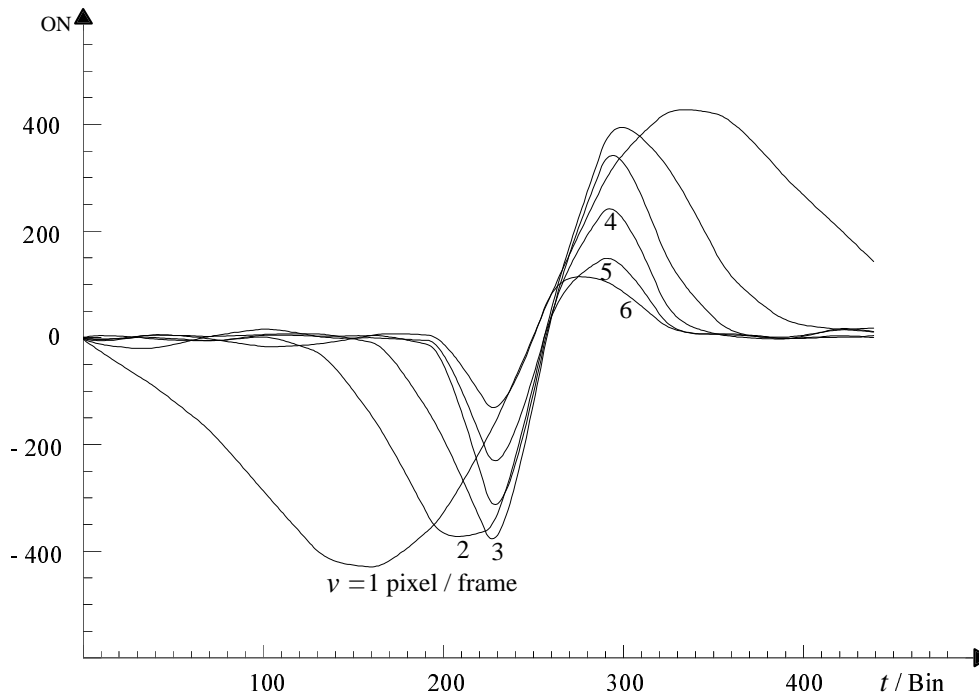


Abb. 3.13: Potentialverlauf einer ON-Zentrum-Bipolarzelle für unterschiedliche Geschwindigkeiten. Die Abbildung zeigt die Abhängigkeit des Bipolarzellenpotentials von der Geschwindigkeit. Je geringer die Geschwindigkeit ist, desto größer ist die Fläche unter der Potentialkurve. Hohe Geschwindigkeiten führen im Vergleich zu geringen Geschwindigkeiten zu kleinen Potentialmaxima. Für die Simulation wurde eine Kante mit den Parametern $g_1=150$ und $g_2=50$ verwendet (Kantenbewegungsrichtung wie in Abbildung 3.10). Die Zahlen an den Kurvenverläufen geben die Geschwindigkeit v der Kante in Pixel/Bild an.

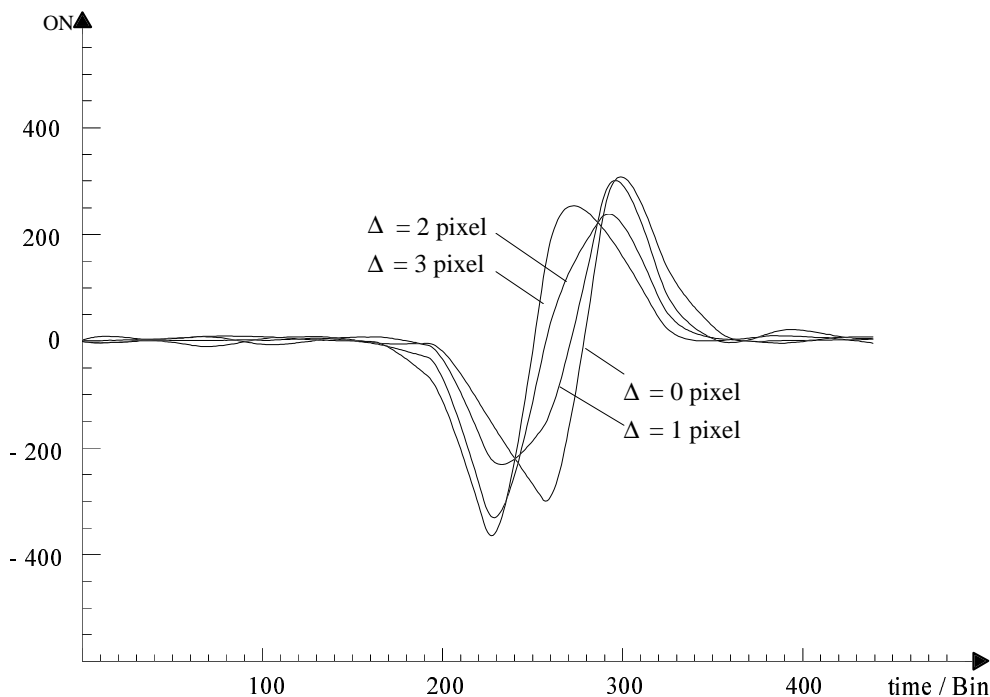


Abb. 3.14: Potentialverlauf einer ON-Zentrum-Bipolarzelle für verschiedene Positionsverschiebungen. Für eine ON-Bipolarzelle sind die Potentiale $B^{\text{ON}}(t)$ für verschiedene Positionsverschiebungen Δ (Kante mit $v = 4$ Pixel/Bild, $g_1=150$ und $g_2=50$) aufgetragen. Die Maxima der Potentiale erreichen unterschiedliche Werte und liegen nicht äquidistant auseinander.

Durch die längere Verweildauer der Kante im rezeptiven Feld sind in der Regel die Maxima für geringe Geschwindigkeiten größer als für hohe. Die Potentialmaxima liegen nicht äquidistant oder geschwindigkeitsabhängig auseinander. Vielmehr spielt die genaue Position der Kante eine entscheidende Rolle.

Dies wird durch die nachfolgende Simulation verdeutlicht. Bei gleichen Parametern der Kante, aber einer festen Geschwindigkeit von $v = 4$ Pixel/Bild und unterschiedlichen Positionsverschiebungen Δ ergeben sich Potentialverläufe, wie sie in Abbildung 3.14 dargestellt sind. Bei den Rezeptorpotentialen führt eine Positionsverschiebung zu einer äquidistanten Verschiebung der Potentialverläufe. Die Potentiale der Bipolarzellen werden hingegen deutlich verändert. Die Steigungen zwischen den Minima und Maxima sind nicht mehr gleich. Eine sich anschließende Transientendetektion durch die Ganglienzellen wird somit auch von der Positionsverschiebung abhängen.

Die retinalen Ganglienzellen besitzen zusätzlich zeitliche Bandpaßeigenschaften und haben im Transientensystem die Aufgabe, lokale Intensitätskontraste in einem großen Dynamikbereich sicher zu detektieren. Kleine Grauwertdifferenzen Δg müssen im Hinblick auf das Signal/Rausch-Verhältnis in einer ausreichend großen Anzahl an Aktionspotentialen kodiert werden, wohingegen große Grauwertdifferenzen mit Rücksicht auf die Übertragungsrate zwischen Bildvorverarbeitung und Hardware-Accelerator keine allzu große Anzahl an Aktionspotentialen erzeugen dürfen. Auch die nachfolgenden Verarbeitungseinheiten, die bilokalen Bewegungsdetektoren, benötigen einen geringen Unterschied in der Aktionspotentialanzahl zwischen kleinen und großen Grauwertdifferenzen. Diese Einschränkung liegt in der additiven Verschaltungsstruktur des Bewegungsdetektors begründet (siehe Kap. 3.2.1). Ein Eingang soll nicht allein den Beitrag zur Überschwelligkeit erbringen. So darf z.B. ein an einem Eingang detektierter hoher Intensitätskontrast nicht zum gleichen Bewegungsdetektorpotential wie eine sich mit der bevorzugten Geschwindigkeit bewegende Kante geringen Intensitätskontrastes, die beide Eingänge stimuliert, führen. Denn dann würden sich nicht bewegende Intensitätskontraste Richtungsdetektoren aktivieren und eine Bewegung in diesem Bildbereich vortäuschen.

Die Abb. 3.15 zeigt an dem Beispiel einer ON-Ganglienzelle, daß die Potentiale der Ganglienzellen bei verschiedenen Grauwertdifferenzen Δg unter diesen Gesichtspunkten ungünstig verlaufen. Die Simulation wurde mit fünf verschiedenen Kanten ($\Delta g = 50, 100, 150, 200, 250$) durchgeführt, die sich mit $v = 4$ Pixel/Bild in Richtung 0° bewegten.

Die Potentialmaxima sind nahezu linear von der Grauwertdifferenz abhängig. Damit auch kleine Grauwertdifferenzen von $\Delta g = 50$ detektiert werden, müßte der Schwellenoffset Θ_0 unterhalb des entsprechenden Potentialmaximums liegen. Die Generierung mehrerer Aktionspotentiale wäre mit Hinblick auf die Eigenschaften des Acceleratorneurons (eine Schellenzeitkonstante) nur unter der Verwendung einer kleinen Schwellenverstärkung V° und einer kleinen Schwellenzeitkonstante τ° möglich. Dies würde aber wiederum zu einer langen Aktionspotentialfolge bei großen Grauwertdifferenzen führen, da das Schwellenpotential nicht schnell genug ansteigen könnte und dadurch zumindest im Bereich des Membranpotentialanstiegs unterhalb dessen liegen würde.

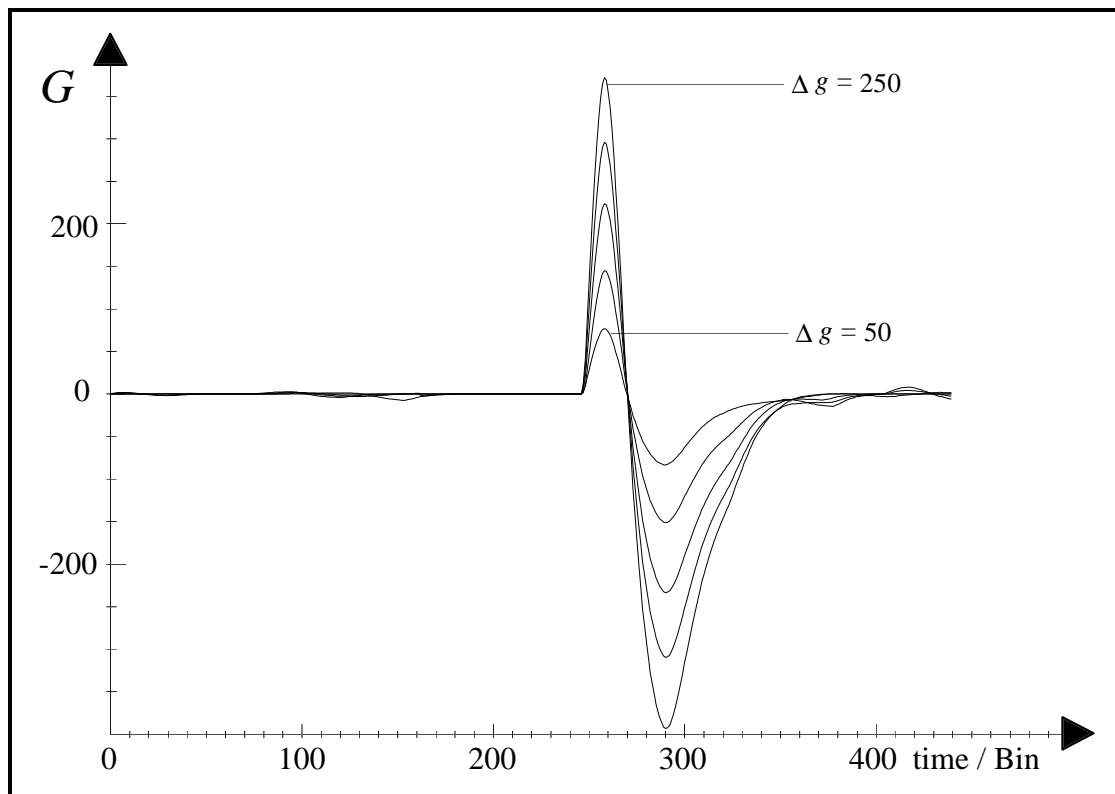


Abb. 3.15: Potentialverläufe einer ON-Ganglienzelle. Die Abbildung zeigt das Potential G einer ON-Ganglienzelle der mittleren Auflösung in Abhängigkeit von der Zeit für fünf verschiedene Kanten ($v = 4$ Pixel/Bild, $\Delta = 0$ Pixel) mit Grauwertdifferenzen $\Delta g = 50, 100, 150, 200$ und 250 . Der Maximalwert des Potentials G hängt linear von der Grauwertdifferenz Δg ab.

Die Abbildung 3.16 zeigt die Potentialmaxima G der Ganglienzellen (**A:** ON, **B:** OFF) für die Geschwindigkeiten v mit $1 \text{ Pixel/Bild} \leq v \leq 6 \text{ Pixel/Bild}$ unter Variation der Grauwertdifferenz Δg . Jede Markierung gibt den Mittelwert der Potentialmaxima für die verschiedenen Positionsverschiebungen Δ an. Das Potentialmaximum einer Positionsverschiebung ist wiederum ein Mittelwert aus zehn Simulationsdurchläufen (Rauschen wie oben beschrieben). Weiterhin ist die doppelte Standardabweichung eingezeichnet.

Da das Potential der ON-Bipolarzellen für diesen Reiz stärker ansteigt als das der OFF-Bipolarzellen sind die Maxima der ON-Ganglienzellen aufgrund der Bandpaßeigenschaften der Ganglienzellen größer als die der OFF-Ganglienzellen. Negative Grauwertdifferenzen durch Vertauschen von g_1 mit g_2 führen zu einem Vertauschen der beiden Antwortverhalten. Die Maxima sind bei gleicher Geschwindigkeit v linear von der Grauwertdifferenz Δg abhängig. Dadurch bedingt werden die absoluten Standardabweichungen mit zunehmender Grauwertdifferenz größer. Die Standardabweichung wächst ebenso mit steigender Geschwindigkeit v , so daß bei höher werdenden Geschwindigkeiten das maximale Potential in Abhängigkeit von der Positionsverschiebung immer mehr schwankt. Die starke Streuung bei großen Kantenschrittweiten folgt aus dem Überspringen der vergleichsweise kleinen RF-Zentren.

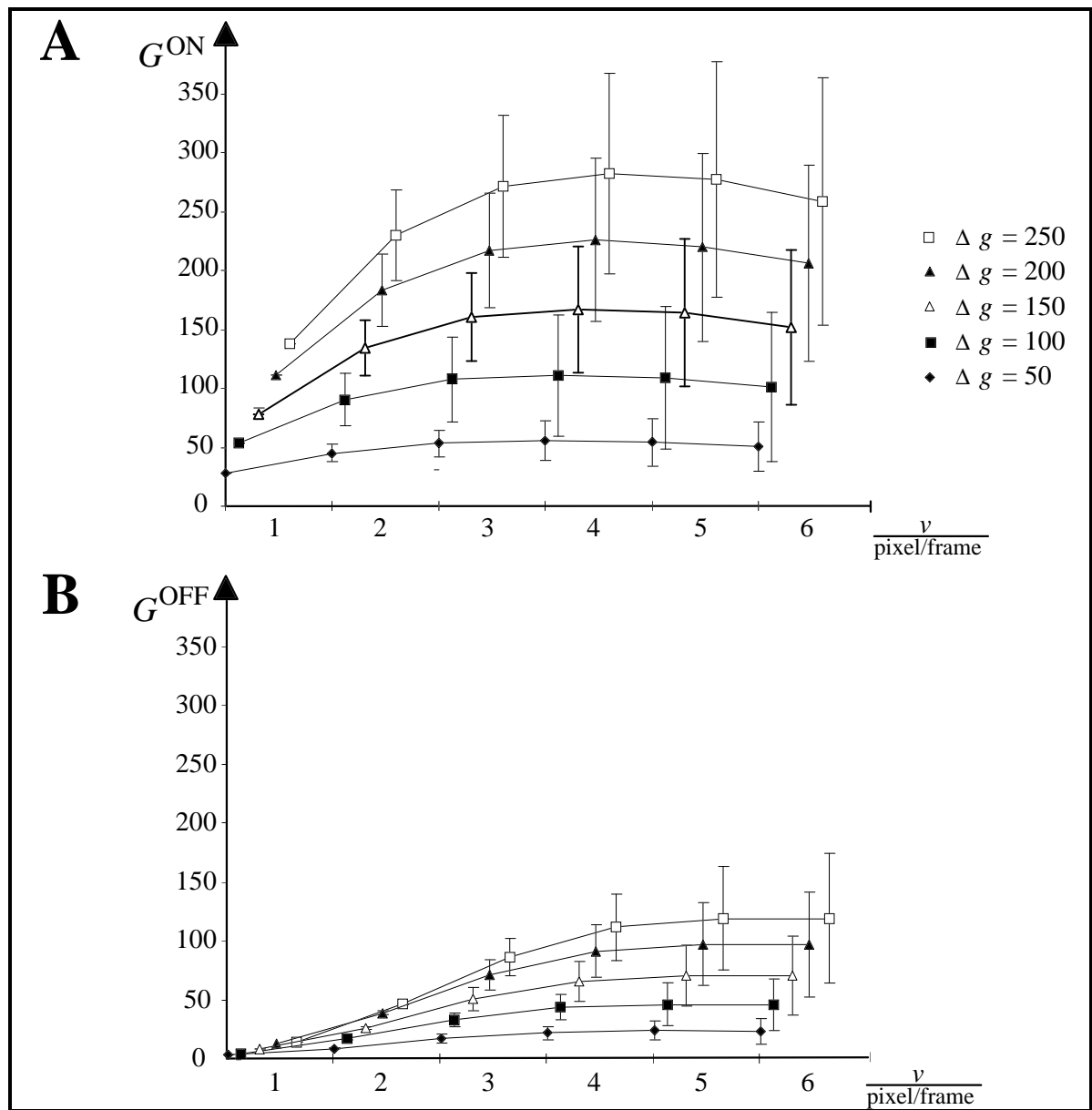


Abb. 3.16: Gemittelte Potentialmaxima G der Ganglienzellen. Die Abbildung zeigt die Mittelwerte der Potentialmaxima G der Ganglienzellen mittlerer Auflösung (**A** ON, **B** OFF) mit doppelter Standardabweichung für verschiedene Geschwindigkeiten v unter Variation der Grauwertdifferenz Δg . Zur besseren Veranschaulichung wurden die Markierungen unterschiedlicher Grauwertdifferenzen leicht versetzt dargestellt und zwischen Markierungen gleicher Grauwertdifferenz Verbindungslinien hinzugefügt.

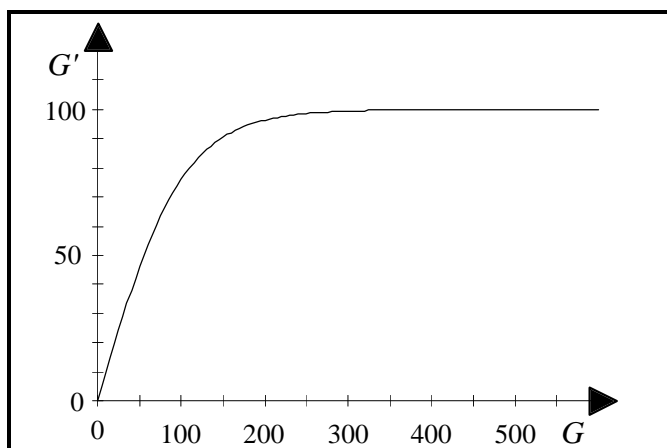


Abb. 3.17: Positiver Ast der Sigmoid-Funktion zur Berechnung des Ganglienzellpotentials. Die Parameter der Sigmoid-Funktion sind $G_{\max}=100$, $\alpha=50$. Dargestellt ist nur der Teil für positive Potentiale. Durch die Verwendung der dargestellten Sigmoid-Funktion wird der Bereich der möglichen Ganglienzellpotentiale eingeschränkt, was zu einer Verringerung der Potentialunterschiede bei Stimulation mit Kanten verschiedener Grauwertdifferenzen führt (vgl. G mit Abb. 3.16).

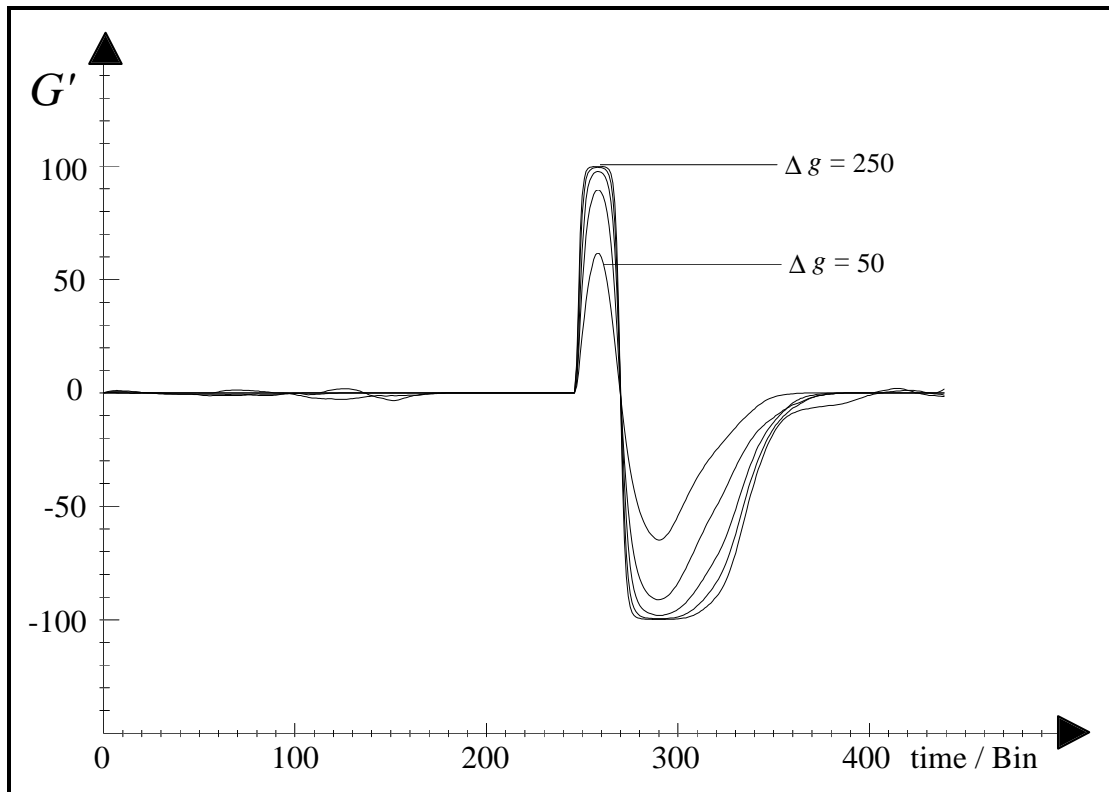


Abb. 3.18: Gestauchte Potentialverläufe einer ON-Ganglienzelle. Die Abbildung zeigt den Potentialverlauf G' einer ON-Ganglienzelle der mittleren Auflösung für fünf verschiedene Kanten ($v = 4$ Pixel/Bild, $\Delta = 0$ Pixel) mit Grauwertdifferenzen $\Delta g = 50, 100, 150, 200$ und 250 . Die Potentialverläufe sind gestaucht. Der Maximalwert des Potentials G hängt nicht mehr linear von der Grauwertdifferenz Δg ab.

Durch die Implementierung einer Sigmoid-Funktion in die Berechnung des Ganglienzellpotentials ist es möglich, schwache Intensitätskontraste unter Berücksichtigung einer nicht zu hohen Aktionspotentialanzahl bei hohen Intensitätskontrasten sicher zu detektieren. Die Parameter $G_{\max} = 100$ und $\alpha = 50$ der Sigmoid-Funktion (Abb. 3.17) sind hierbei so gewählt, daß kleine Potentiale nahezu unverändert bleiben und große stark gestaucht werden. Dadurch wird eine Anpassung der Schwellenparameter an kleine Potentialwerte möglich. Die Implementierung kann ohne großen Zuwachs an Rechenzeit durch eine Tabelle erfolgen.

Die Auswirkung der Sigmoid-Funktion auf den Potentialverlauf zeigt die Abbildung 3.18. Es wurden die gleichen Netzwerkparameter wie zuvor benutzt. Das Potential G wurde lediglich über die Sigmoid-Funktion in das Potential G' umgerechnet. Die Potentialmaxima sind nun nicht mehr linear von der Grauwertdifferenz abhängig. Es kommt zu einer immer stärkeren Angleichung der Potentialverläufe für größer werdende Grauwertdifferenzen. Das Maximum des Potentialverlaufs für $\Delta g = 250$ ist nur etwa um 60% größer als das des Potentialverlaufs für $\Delta g = 50$. Dies ist im Hinblick auf die geringe Übertragungskapazität von der Vorverarbeitung auf den Accelerator vorteilhaft.

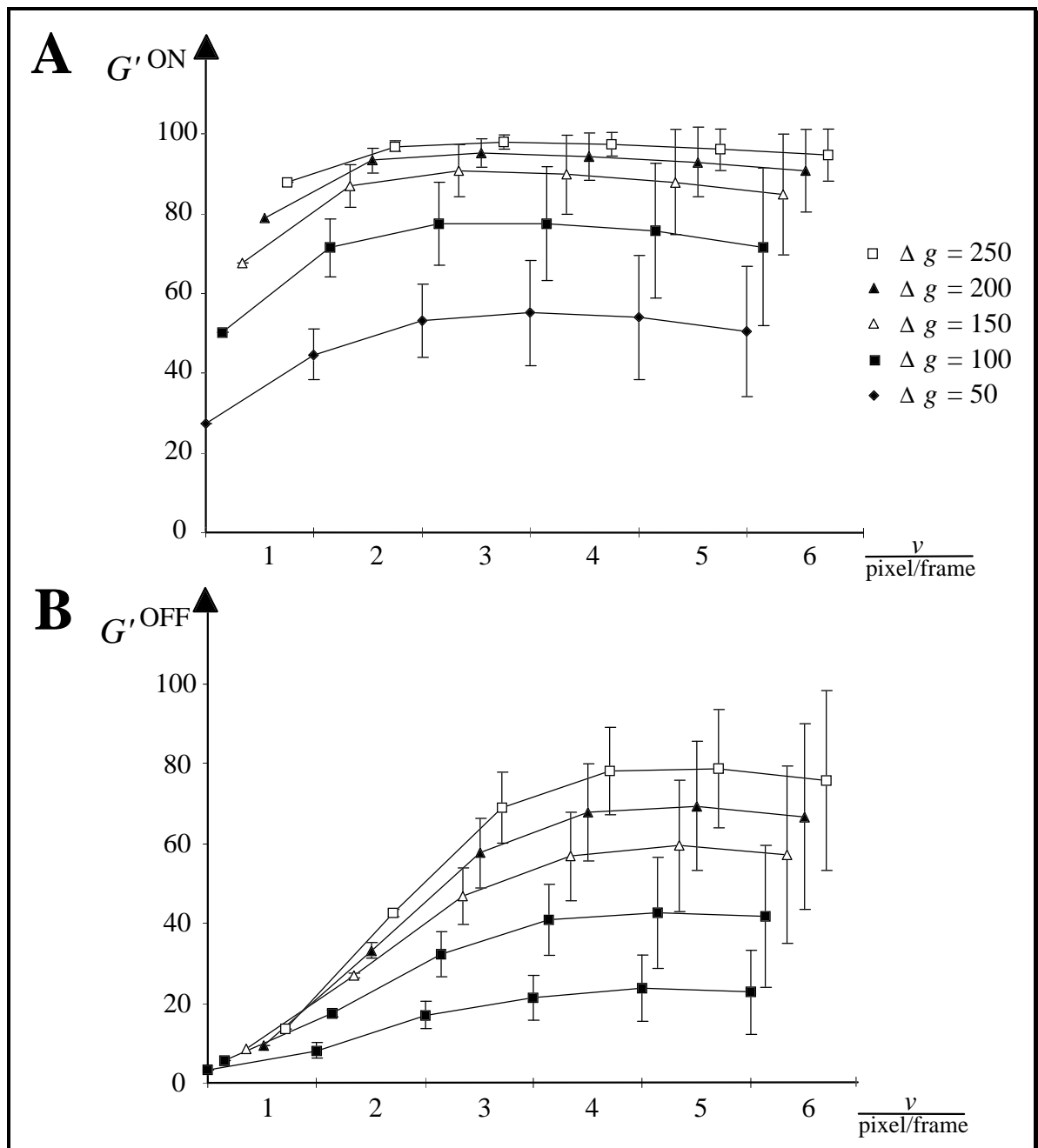


Abb. 3.19: Gemittelte Potentialmaxima G' der Ganglienzellen. Die Abbildung zeigt die Mittelwerte der Potentialmaxima G' der Ganglienzellen mittlerer Auflösung (**A**: ON, **B**: OFF) mit doppelter Standardabweichung für verschiedene Geschwindigkeiten v unter Variation der Grauwertdifferenz Δg . Die Simulationen wurden mit Sigmoid-Funktion durchgeführt. Nähere Erläuterungen im Text.

Der Unterschied zwischen den Potentialen G und G' wird aus dem Vergleich der beiden Abbildungen 3.16 und 3.19 deutlich. Mit implementierter Sigmoid-Funktion sind die Potentialmaxima der ON-Ganglienzellen weiterhin größer als die der OFF-Ganglienzellen. Die Unterschiede zwischen den Kurven nehmen für größer werdende Grauwertdifferenzen ab. Die Streuung, die aus der unterschiedlichen Positionsverschiebung resultiert, wird für große Grauwertdifferenzen deutlich geringer.

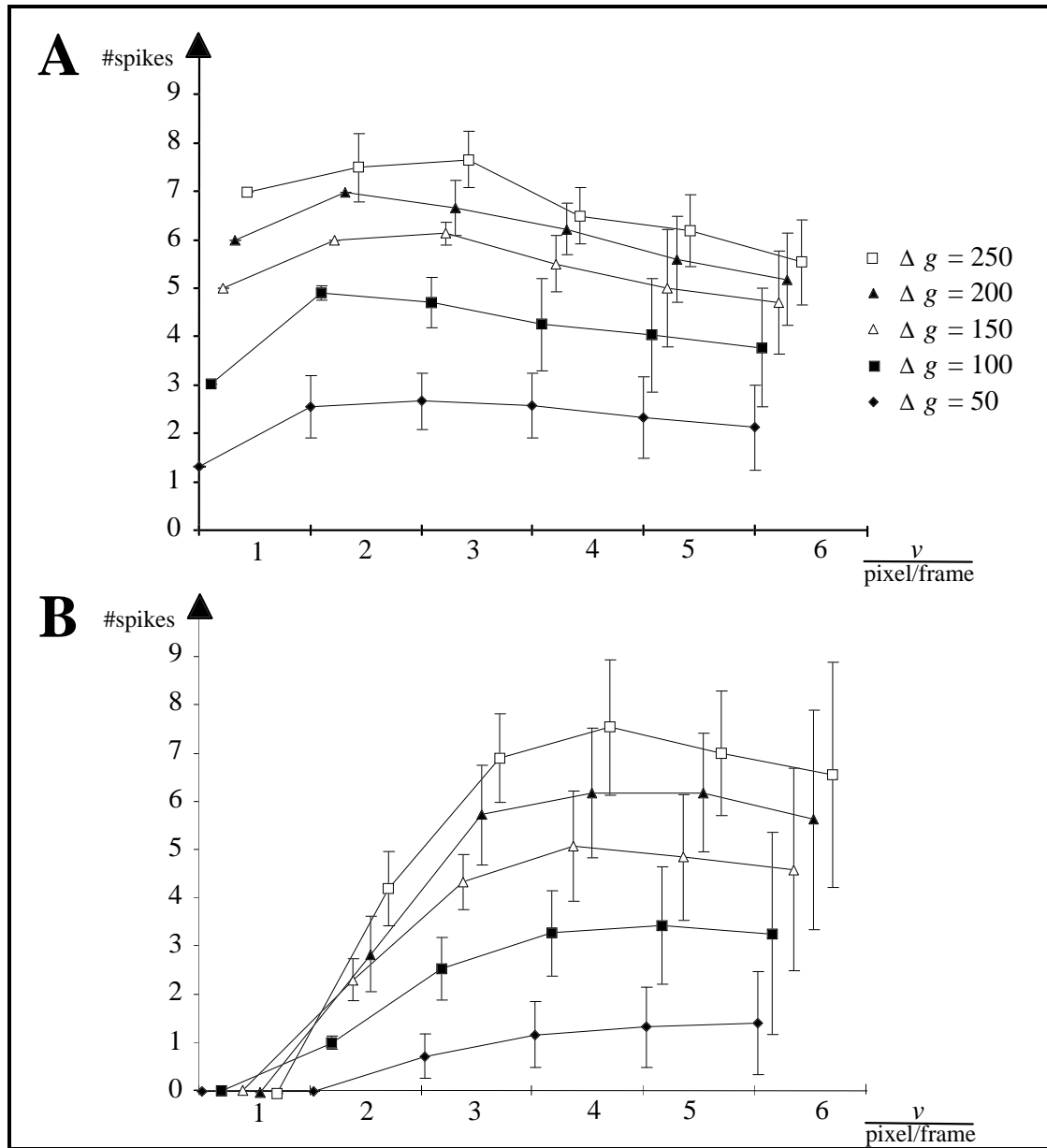


Abb. 3.20: Mittlere Anzahl der von Ganglienzellen generierten Aktionspotentiale. Die Abbildung zeigt die mittlere Anzahl von ON- (A) und OFF- (B) Ganglienzellen der mittleren Auflösung generierten Aktionspotentiale pro Kantendurchlauf mit doppelter Standardabweichung für verschiedene Geschwindigkeiten v unter Variation der Grauwertdifferenz Δg an. Die Simulationen wurden unter Verwendung der Sigmoid-Funktion durchgeführt. Nähere Erläuterungen im Text.

Die Anzahl der generierten Aktionspotentiale ist in Abbildung 3.20 dargestellt. Jede Markierung ergibt sich wie oben dargestellt als Mittelwert aus Simulationen mit unterschiedlicher Positionsverschiebung, die jeweils zehnmal mit Rauschen durchgeführt wurden. Die Aktionspotentialanzahl wird wie das Potentialmaximum in erster Linie durch die Grauwertdifferenz Δg bestimmt. ON-Ganglienzellen generieren für alle simulierten Kantengeschwindigkeiten Aktionspotentiale, wohingegen OFF-Ganglienzellen bei der geringsten Geschwindigkeit $v = 1$ Pixel/Bild keine Aktionspotentiale generieren. Große Geschwindigkeiten ($v \geq 4$ Pixel/Bild) lassen die OFF-Ganglienzellen ungefähr ebensoviele Aktionspotentiale generieren wie die ON-Ganglienzellen. Die Maxima liegen abhängig von

der jeweiligen Grauwertdifferenz bei ON-Ganglienzellen bei einer Geschwindigkeit von $v = 2$ Pixel/Bild oder $v = 3$ Pixel/Bild, die der OFF-Ganglienzellen bei $v = 4$ Pixel/Bild oder $v = 5$ Pixel/Bild. Die Standard-abweichungen sind bei den Simulationen der OFF-Ganglienzellen größer als die der ON-Ganglienzellen, da die Aktionspotentialfolgen bei den OFF-Ganglienzellen länger sind als bei den ON-Ganglienzellen. Aus den Ganglienzellaktivitäten können im allgemeinen keine Rückschlüsse auf die Geschwindigkeit gezogen werden.

Zur Untersuchung der Geschwindigkeitscharakteristik retinaler Ganglienzellen bezüglich „Kantenschärfe“ wurde die Kante durch eine Welle einer sinusförmigen Intensitätsverteilung (Maximalwert 200, Minimalwert 100, siehe Abbildung 3.21 A) ersetzt. Die Welle bewegte sich mit unterschiedlicher Geschwindigkeit in Richtung 180° . Die Eingabebilder waren bei der Untersuchung der feinen und mittleren Auflösung $128 \text{ Pixel} \times 128 \text{ Pixel}$ groß, bei der großen $256 \text{ Pixel} \times 256 \text{ Pixel}$. Eine Sequenz bestand aus acht Bildern. Die Welle wurde so bewegt, daß sie im vierten Bild der Sequenz in der Bildmitte lag (Positionsverschiebung $\Delta = 0$ Pixel).

In Abbildung 3.22 sind die Simulationsergebnisse für die drei Auflösungen (A: fein, B: mittel und C: grob) dargestellt. Es ist jeweils die mittlere Anzahl an Aktionspotentialen einer OFF-Ganglienzelle als Grauwert in Abhängigkeit von der Wellenlänge und der Geschwindigkeit aufgetragen. Hierbei werden sowohl die räumlichen als auch die zeitlichen Bandpaßeigenschaften der retinalen Ganglienzellen deutlich. Die Ganglienzellen der feinen Auflösung antworten am stärksten bei Geschwindigkeiten von 1 Pixel/Bild bis 3 Pixel/Bild, die der mittleren Auflösung von etwa 2 Pixel/Bild bis 7 Pixel/Bild und die der groben Auflösung von etwa 2 Pixel/Bild bis 13 Pixel/Bild.

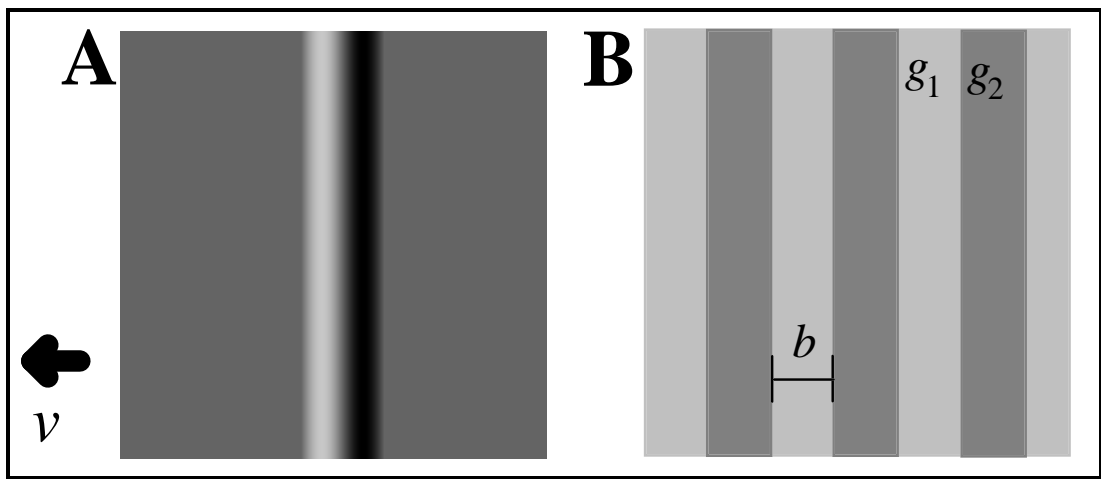


Abb. 3.21: Stimuli zur Untersuchung der zeitlichen und räumlichen Übertragungseigenschaften der Ganglienzellen. Die Abbildung zeigt in **A** eine Welle einer sinusförmigen Intensitätsverteilung. Der homogene Hintergrund hat den Grauwert 100. Der Maximalwert beträgt 200, der Minimalwert entsprechend 0. Die Welle bewegt sich mit der Geschwindigkeit v in Richtung 180° fort. Der Stimulus enthält einen Grauwertabfall, der bevorzugt von den OFF-Zellen verarbeitet wird. In **B** ist der Gitterreiz dargestellt. Die Breite der einzelnen Gitterbalken ist b , die Grauwerte g_1 und g_2 . Der Gitterreiz wird ebenfalls mit einer Geschwindigkeit v in Richtung 180° bewegt.

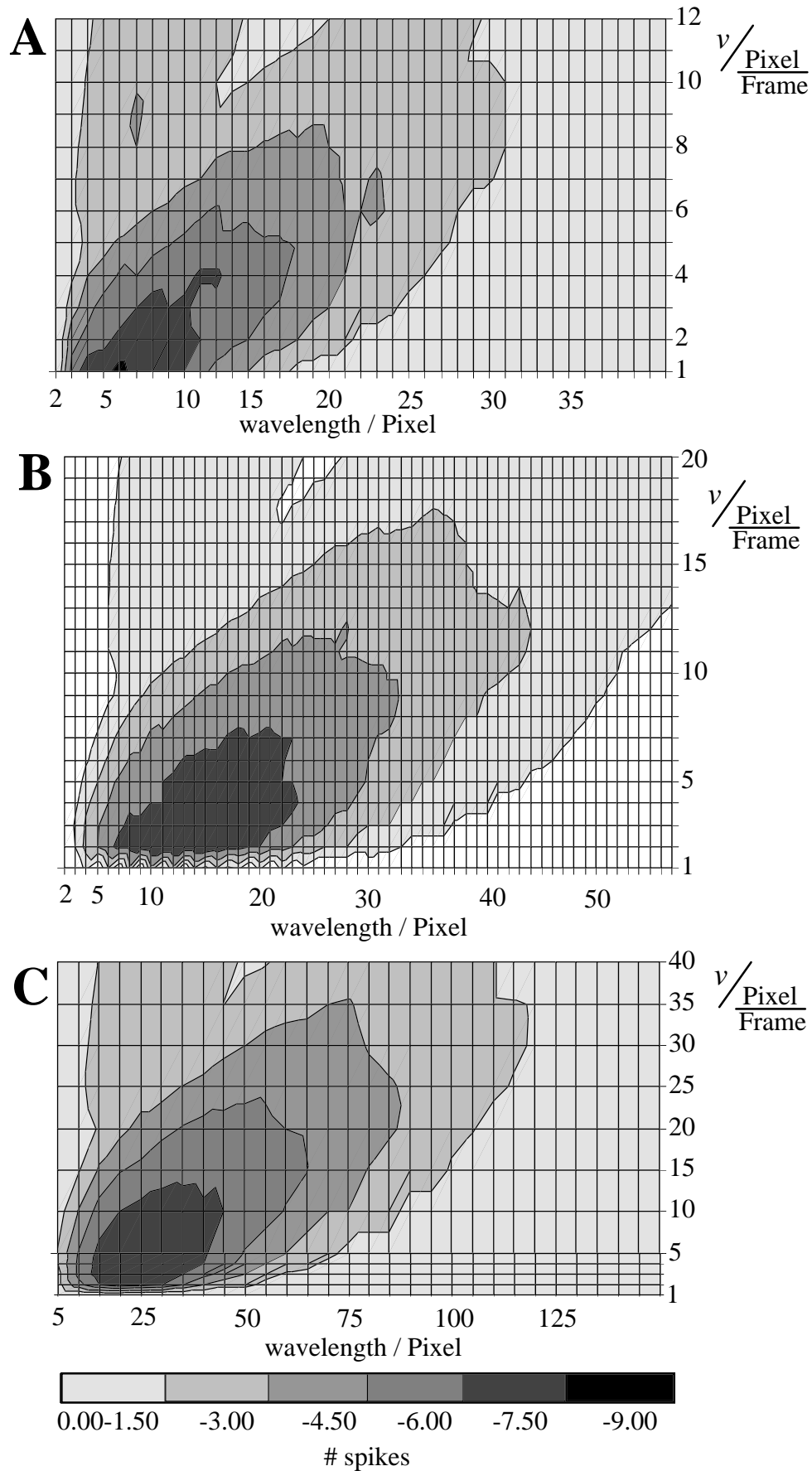


Abb.3.22: Geschwindigkeitscharakteristik retinaler Ganglienzellen bezüglich „Kantenschärfe“. **A:** Auflösung mit Abtastweite $d = 2$ Pixel. **B:** Abtastweite $d = 4$ Pixel **C:** Abtastweite $d = 8$ Pixel. Erläuterungen im Text. (Aus Gründen der besseren Darstellung wurden die Achsen in A, B und C unterschiedlich skaliert.)

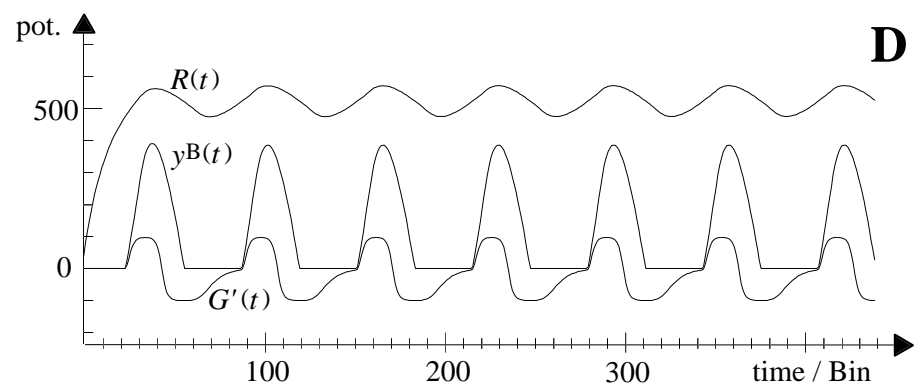
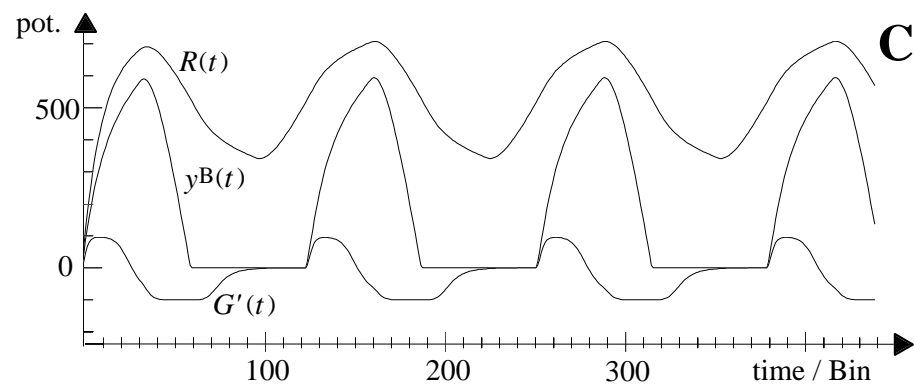
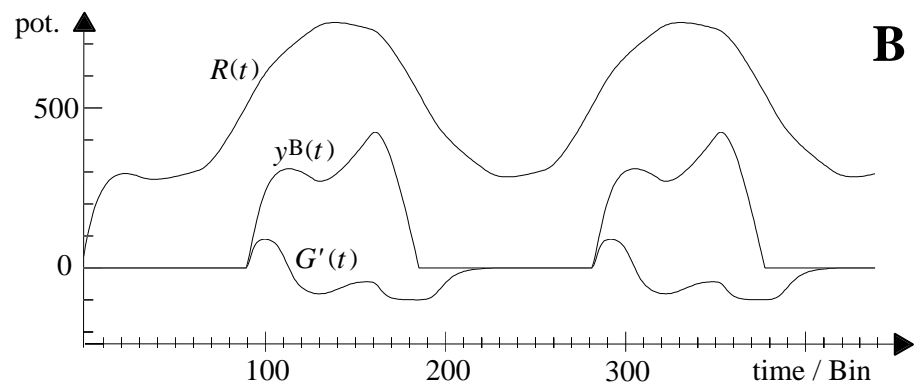
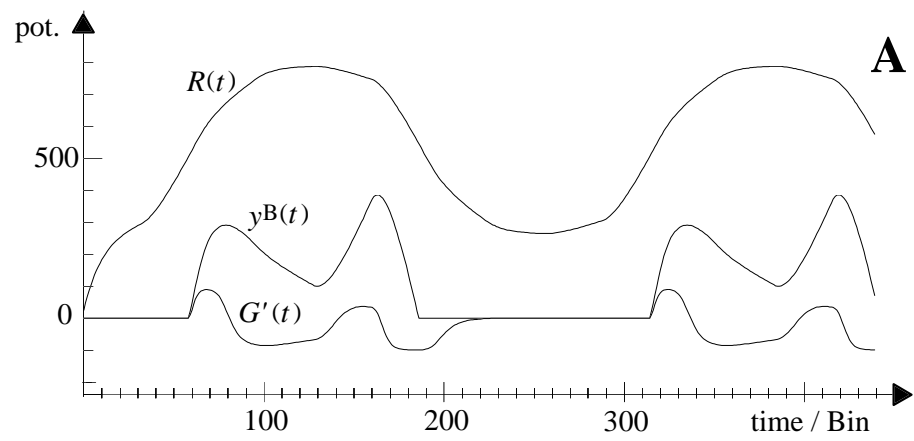


Abb. 3.23: Rezeptor-, Bipolarzellen- und Ganglienzellenpotential in Abhängigkeit von der Balkenbreite des Stimulationsgitters. In den Abbildungen **A**, **B**, **C** und **D** sind jeweils Rezeptorpotential $R(t)$, ON-Bipolarzellenpotential $y^B(t)$ und ON-Ganglienzellenpotential $G'(t)$ für die Gitterbalkenbreiten $b=16, 12, 8$ und 4 Pixel ($g_1=150, g_2=50$) gegen die Zeit aufgetragen (Stimulus aus Abb. 3.21B). Die Ganglienzellen sind aufgrund kurzer Zeitkonstanten an ihren Eingängen in der Lage, auch schnell aufeinanderfolgende Balkenreize zu detektieren (**D**) .

Die Aktivität ist zusätzlich von der Wellenlänge abhängig. Die stärksten Antworten gibt es für Ganglienzellen der feinen Auflösung bei einer Wellenlänge von 4 Pixel bis 10 Pixel, für die mittlere Auflösung bei einer Wellenlänge von 7 Pixel bis 22 Pixel und für die grobe Auflösung bei einer Wellenlänge von 15 Pixel bis 45 Pixel. Damit sind sowohl ein großer Geschwindigkeitsbereich als auch ein großer Wellenlängenbereich ausreichend abgedeckt.

In der nächsten Simulation (Abbildung 3.23) wird die Aktivierung der Verarbeitungseinheiten der mittleren Auflösung im ON-Pfad auf einen Stimulus in Form eines Gitterreizes (Abbildung 3.21 B) mit den Grauwerten $g_1=150$, $g_2=50$, den Balkenbreiten $b = 4, 8, 12$ und 16 Pixel und einer Geschwindigkeit von $v = 4$ Pixel/Bild ($\Delta=0$ Pixel) in Richtung 0° untersucht.

Das Rezeptorpotential $R(t)$ pendelt in Abhängigkeit von Grauwert und Balkenbreite nach dem Anschalteffekt ($t < 50$ Bin) zwischen Maximal- und Minimalwert. Die Differenz zwischen Maximal- und Minimalwert wird aufgrund der zeitlichen Tiefpaßfilterung ($\tau=10$ Bin) mit kleiner werdender Balkenbreite immer geringer.

Entsprechend der räumlichen Bandpaßfiltereigenschaften der Bipolarzellen und der Gittergeschwindigkeit können für $b = 12$ Pixel und $b = 16$ Pixel noch Vorder- und Hinterkante eines Balkens detektiert werden (Abb. 3.23 A, B). Für kleinere Balkenbreiten ist dies nicht mehr möglich (Abb. 3.23 C, D).

Der Potentialverlauf der Ganglienzelle in Abbildung 3.23 D zeigt, daß auch rasch aufeinanderfolgende, dünne Balken zu einem positiven Ganglienzellenpotential G' führen. Wären die Zeitkonstanten der Ganglienzelleneingänge zu groß gewählt worden, wäre dies nicht mehr der Fall.

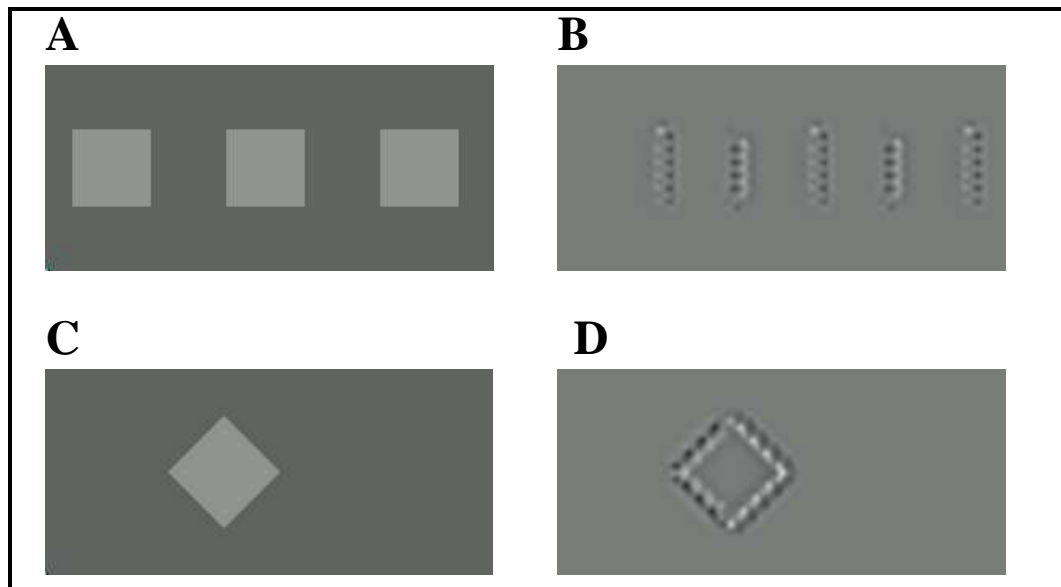


Abb. 3.24: Vorverarbeitung einfacher geometrischer Objekte. Links (A, C) sind Bilder zweier verschiedener Eingabesequenzen dargestellt. Die hellgrauen Objekte bewegen sich von links nach rechts. Auf der rechten Seite (B, D) sind die Simulationsergebnisse als Rückprojektionen (Kopecz et al. 1996) der rezeptiven Felder aktiver Ganglienzellen dargestellt. Das Simulationsergebnis (B) zeigt, daß die transienten Ganglienzellen nicht von quer zur Bewegung liegenden Kontrastkanten aktiviert werden.

Bevor die Vorverarbeitung auf reale Eingabebilder angewendet wird, sollen zunächst die wichtigste Eigenschaft der Vorverarbeitung und die Darstellung der Simulationsergebnisse anhand konstruierter Eingabebilder mit einfachen geometrischen Objekten erläutert werden.

Die Aufgabe des Transientensystems ist die Extraktion und Analyse von Veränderungen in der Eingabesequenz, damit diese für die Szenensegmentierung genutzt werden können. Die erste Verarbeitungsstufe dieses Teilsystems bilden die modellierten retinalen Ganglienzellen. Sie generieren ausschließlich Aktionspotentiale, wenn sich ihr Eingangssignal signifikant verändert. Dies ist insbesondere der Fall, wenn Kanten ihr rezeptives Feld durchqueren.

Die Abbildung 3.24 A zeigt hellgraue Vierecke ($g_1 = 150$), die sich mit einer Geschwindigkeit von $v = 4$ Pixel/Frame vor einem dunkelgrauen Hintergrund ($g_2 = 50$) in Richtung 0° bewegen. Die Simulationsparameter entsprechen denen aus Tabelle 3.1. In Abbildung 3.24 B ist das Simulationsergebnis der Vorverarbeitung mit mittlerer Auflösung als Rückprojektionen der rezeptiven Felder aktiver Ganglienzellen dargestellt (Kopecz et al. 1996). Generiert eine Ganglienzelle ein Aktionspotential, so wird dessen Sensitivitätsprofil (siehe Abbildung 3.4) in den Bildbereich projiziert. ON-Bereiche des rezeptiven Feldes werden hellgrau, OFF-Bereiche dunkelgrau dargestellt. Zur besseren Veranschaulichung werden die rezeptiven Felder durch eine zeitliche Tiefpaßfilterung ($\tau = 20$ Bin) kumuliert dargestellt.

Obwohl die Vierecke aus Abbildung 3.24 A und C die gleiche Grauwertdifferenz zum Hintergrund, die gleiche Geschwindigkeit und die gleiche Form aufweisen, sind die Simulationsergebnisse unterschiedlich. In Abbildung 3.24 B werden die Kanten, welche parallel zur Bewegungsrichtung liegen, nicht detektiert. Für die lokalen Transientendetektoren sind diese Kanten über einen zu großen Zeitraum konstant, wodurch sie von ihnen nicht detektiert werden. Diese Kanten können durch die X-Zellen des Kontur/Form-Systems detektiert werden (z.B. Weitzel 1998), da die scheinbare Stationarität der Kante in den rezeptiven Feldern der X-Zellen ausreicht, um ein anhaltendes Potential aufzubauen. Nicht parallel zur Bewegungsrichtung ausgerichtete Kanten werden hingegen von den Y-Ganglienzellen detektiert, wie in Abbildung 3.24 D zu sehen ist.

3.1.2.2 Simulationen mit realen Szenen

Nachdem das Antwortverhalten der Ganglienzellen auf künstliche Stimuli dargestellt wurde, sollen in diesem Kapitel Sequenzen mit realen Objekten als Stimuli benutzt werden.

Die Abbildung 3.25 zeigt die vom FhG-IITB Karlsruhe bereitgestellte Szene „Durlacher Tor“ zu zwei verschiedenen Zeitpunkten ($t = 5$ Bin (A), $t = 420$ Bin (B)). Die Szene beinhaltet eine von einem Hochhaus aufgenommene Straßenkreuzung mit fahrenden Kraftfahrzeugen. Die Kraftfahrzeuge sind aufgrund der großen Entfernung zur Kamera klein und bewegen sich im Bild nur mit kleinen Geschwindigkeiten von maximal $v = 2$ Pixel/Bild. Die in den Simulationen verwendeten Modellneuronenparameter entsprechen denen in Tabelle 3.2 aufgelisteten.

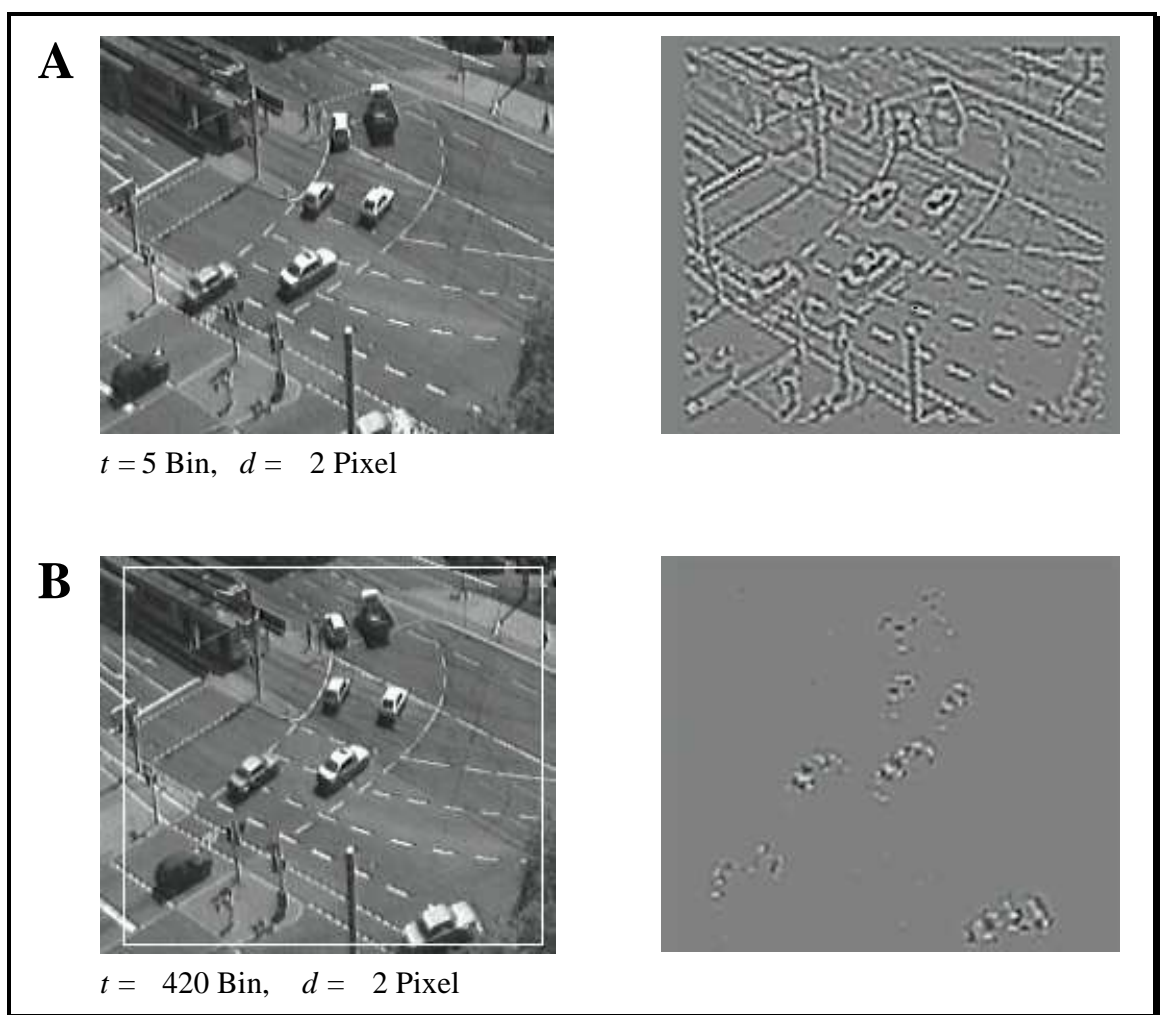


Abb. 3.25: Sequenz „Durlacher Tor“ vollständig mit Ganglienzellen der feinen Auflösung vorverarbeitet. Die Abbildung zeigt jeweils links das Originalbild und den Abtastbereich (weißer Rahmen in B) sowie rechts die rückprojizierten rezeptiven Felder der aktiven Ganglienzellen. Die Sequenz wurde von der FhG-IITB-Karlsruhe zur Verfügung gestellt (Gengenbach und Nagel 1994). Die Rückprojektion (Kopocz et al. 1996) der rezeptiven Felder aktiver Ganglienzellen liefert hellgraue Markierungen für die ON-Bereiche und dunkelgraue für die OFF-Bereiche (zeitliche Tiefpaßfilterung mit $\tau = 20$ Bin). Die Simulation wurde mit 100×90 Abtastpunkten in der feinen Auflösung ($d = 2$ Pixel) durchgeführt. Im Simulationsergebnis (A) sind aufgrund des Anschalteffektes alle Kanten zu sehen, in (B) sind die ortsfesten Kanten nicht mehr enthalten.

In der Abbildung sind links die Eingabebilder und der Abtastbereich zu sehen. Der Abtastbereich ist der Bereich, in dem alle RF-Mittelpunkte der Ganglienzellen liegen. Er ist durch ein weißes Rechteck markiert (B). Rechts daneben sind die Rückprojektionen der rezeptiven Felder aktiver Ganglienzellen dargestellt.

Nahezu das gesamte Eingabebild wurde mit den Verarbeitungseinheiten der feinen Auflösung (Abtastweite $d = 2$ Pixel) abgetastet. Der Anschalteffekt (bei $t = 5$ Bin), der z.B. auch aus einer schnellen Kamerabewegung resultieren könnte, ist in Abbildung 3.25 A zu sehen. Alle Kanten - auch die Kanten der sich nicht bewegenden Objekte - werden von den transienten Ganglienzellen detektiert. Durch die Unterabtastung mit Tiefpaßfilterung wirkt das Bild verschmiert.

Nach dem Anschalteffekt werden in diesem Beispiel nur noch durch Fahrzeugbewegungen verursachte zeitliche Veränderungen extrahiert. Die Abbildung 3.25 zeigt dies in B. Die Bewegung aller acht Kraftfahrzeuge wird von den Ganglienzellen detektiert. Die Antworten der Ganglienzellen auf die beiden Fahrzeuge im oberen Bereich des Bildes sind nicht so stark, da sie sich mit einer geringen Geschwindigkeit bewegen. Die Bewegung des dunklen Fahrzeugs im linken unteren Bildbereich führt aufgrund seines kleinen Intensitätskontrastes zur Straße ebenfalls zu einer geringen Aktivität. Eine genaue Klassifizierung der Fahrzeuge würde nicht möglich sein, da zum einen die Fahrzeuge im Bild zu klein sind und zum anderen nicht alle Fahrzeugteile eine lokal detektierbare Bewegung vollziehen.

Für die weitere Verarbeitung sind zwei Eigenschaften der Vorverarbeitung von Bedeutung. Zum einen darf aus technischer Sicht die Gesamtaktivität der Ganglienzellen nicht zu groß sein, damit die Übertragung der Vorverarbeitungsdaten auf ein Hardwaresystem auf die Rechenzeit nicht bremsend wirkt, zum anderen soll dem sich an die Vorverarbeitung anschließenden neuronalen Netzwerk möglichst kein Takt durch die Abtastung der Kamera aufgeprägt werden.

Der Zeitverlauf der neuronalen Aktivität läßt sich durch die *Multiple Unit Activity* (MUA) beschreiben. Sie ist hier als die Summe aller Modellneuronausgaben y_i einer bestimmten Gruppe M , wie z.B. alle Modellneuronen einer Schicht, definiert.

$$\text{MUA}(t) = \sum_{i \in M} y_i(t) \quad (\text{Gl. 3.18})$$

Die MUA der ON- und OFF-Ganglienzellen sind für die oben beschriebene Simulation in Abbildung 3.26 aufgetragen. Eine Verteilung, die ausschließlich Ganglienzellenaktivität bei der Darbietung eines neuen Einzelbildes aufweist, liegt nicht vor. Ohne den Anschalteffekt ($t < 50$ Bin) generieren alle ON- und OFF-Ganglienzellen der Vorverarbeitung in 1535 Simulationsschritten 19438 bzw. 18020 Aktionspotentiale, was im Mittel 25 Aktionspotentialen pro Zeitschritt entspricht. Die Gesamtaktivität, die natürlich von der Eingabesequenz abhängt, liegt in diesem Beispiel deutlich unter einem Prozent.

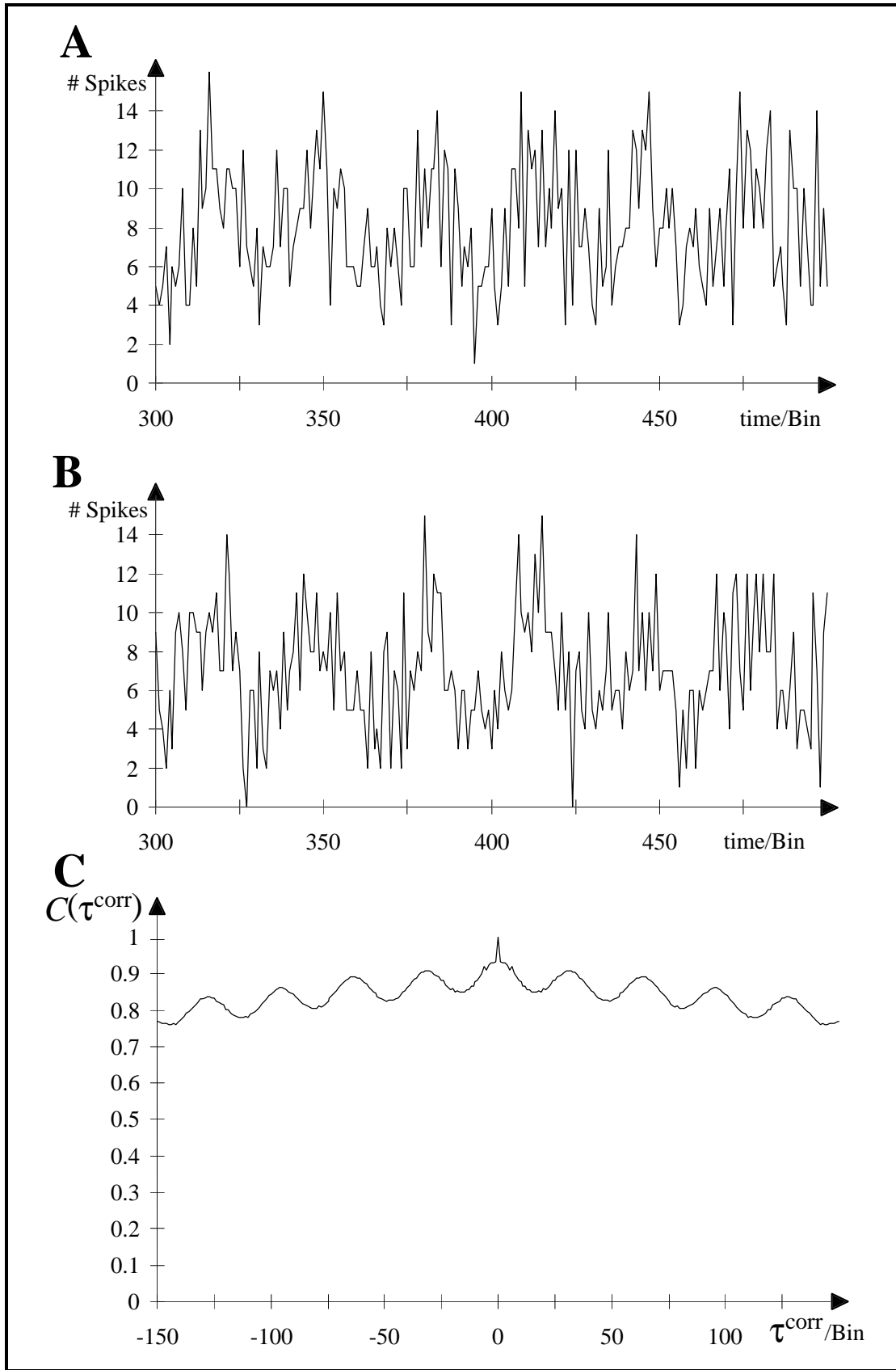


Abb. 3.26: Aktivität der Ganglienzellen (Sequenz „Durlacher Tor“). Dargestellt sind in (A) und (B) jeweils ein zeitlicher Ausschnitt der MUA (Gl. 3.18) aller ON- bzw. OFF-Ganglienzellen der feinen Auflösung. Verarbeitet wurde die Sequenz „Durlacher Tor“. Das Abstraster setzte sich aus 100×90 Abtastpunkten zusammen. In (C) ist die Autokorrelationsfunktion (Gl. 3.19) der aus ON- und OFF-Ganglienzellen resultierenden Gesamtaktivität aufgetragen.

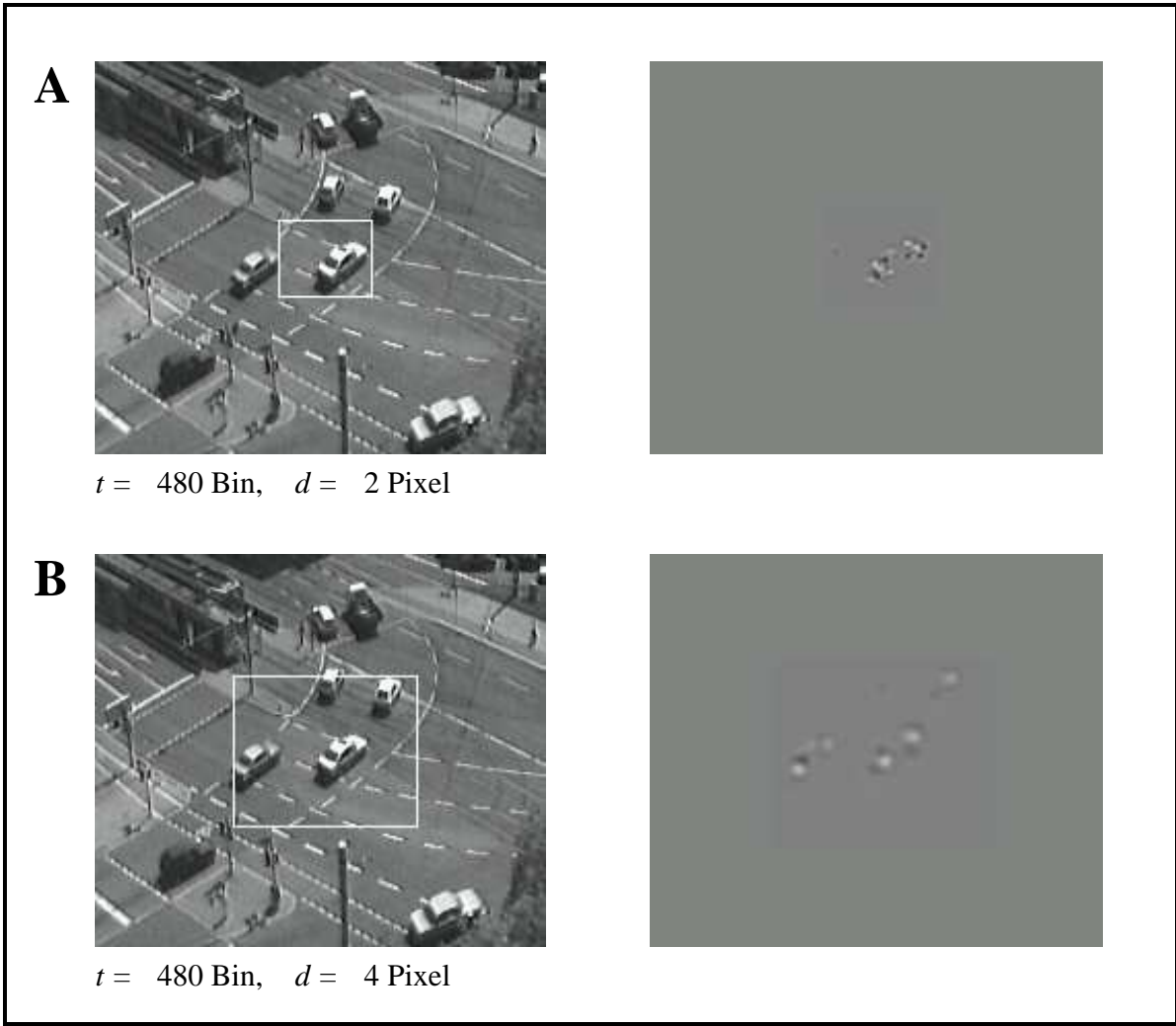


Abb. 3.27: Vorverarbeitung der Sequenz „Durlacher Tor“. Die Abbildung zeigt jeweils links das Originalbild mit Abtastbereich (weißer Rahmen) und rechts die rückprojizierten rezeptiven Felder der aktiven Y-Zellen. In **A** und **B** ist der Abtastbereich - im Sinne des in Kapitel 3.1 beschriebenen Retinamodells - der feinen ($d = 2 \text{ Pixel}$) bzw. mittleren ($d = 4 \text{ Pixel}$) Auflösung räumlich eingeschränkt (29×16 Abtastpunkte). Das Simulationsergebnis der groben Auflösung ($d = 8 \text{ Pixel}$) ist nicht dargestellt.

Ob dem System ein Takt von der Vorverarbeitung aufgeprägt wird, läßt sich mit der Autokorrelationsfunktion $C(\tau^{\text{corr}})$ für eine MUA zwischen den Simulationszeiten t_1 und t_2 untersuchen.

$$C(\tau^{\text{corr}}) = \frac{\sum_{t_1}^{t_2} MUA(t) \cdot MUA(t - \tau^{\text{corr}})}{\sum_{t_1}^{t_2} MUA(t) \cdot MUA(t)} \quad (\text{Gl. 3.19})$$

In Abbildung 3.26 (C) ist die Autokorrelationsfunktion für $t_1 = 50 \text{ Bin}$ und $t_2 = 1375 \text{ Bin}$ aufgetragen. Der Verlauf der Korrelationsfunktion zeigt lokale Maxima ungefähr im zeitlichen Abstand zweier aufeinanderfolgender Kamerabilder. Diese sind aber unter Beachtung der Differenz zu den lokalen Minima nicht besonders ausgeprägt. Somit wird dem neuronalen Netz von der Vorverarbeitung kein dominanter Takt aufgezwungen.

Eine Verarbeitung des gesamten Bildbereiches mit Ganglienzellen der feinen Auflösung würde zu einer zu großen Anzahl an Modellneuronen führen. Deswegen wird, wie in Kapitel 3.1.1.2 beschrieben, die Abtastung mit hoher Auflösung auf einen kleinen Bereich beschränkt. In Abbildung 3.27 A ist dies für die feine und in Abbildung 3.27 B für die mittlere ($d = 4$ Pixel) Auflösung dargestellt. Die mittlere Auflösung liefert hierbei keine Objektkonturen mehr, da die RF-Zentren bereits in der Größenordnung der Fahrzeugdächer und -motorhauben liegen. Die grobe Auflösung liefert nur vereinzelt Aktivität und kann bei dieser Sequenz nicht genutzt werden. Das System könnte daher mit der feinen Auflösung ein Objekt zur genauen Analyse verfolgen und die mittlere und grobe Auflösung zur Aufmerksamkeitssteuerung benutzen (z.B. Pauly et al. 1997).

Die Vorteile des Vorverarbeitungssystems mit unterschiedlichen Auflösungen zeigen sich auch bei der Verarbeitung der Sequenz „Gerade“ (Abb. 3.28). Die Sequenz zeigt eine typische Autobahnfahrt aus der Sicht eines sich auf der Autobahn bewegendes Fahrzeugs. Die Kamera ist zunächst nach vorne gerichtet. Ein Kontur/Form-System könnte in diesem Fall das vorausfahrende Fahrzeug analysieren, da dieses nur eine sehr geringe Relativbewegung zur Kamera aufweist. Die feine Auflösung des Transientensystems detektiert lediglich den Mittelstreifen der Autobahn und Relativbewegungen, die z.B. durch Bodenwellen usw. hervorgerufen werden (Abb. 3.28 A). Die Bewegung eines überholenden Fahrzeugs wird hingegen von der groben Auflösung des Transientensystems detektiert (Abb. 3.28 B). Die Auflösung ist allerdings zu grob, um Details detektieren zu können.

In Abbildung 3.28 C und D sind mögliche Kamerablickpositionen durch Verschieben der Abtastbereiche dargestellt. Es wird hierbei deutlich, daß das Transientensystem nicht die vollständige Kontur eines Objektes extrahieren kann. Kanten, die sich parallel zu ihrer Orientierung bewegen, werden vom Transientensystem nicht detektiert. Lediglich die Enden einer Kante führen zu einer transienten Antwort der Ganglienzellen. In der Sequenz „Gerade“ wird dies besonders an den Mittelstreifen deutlich.

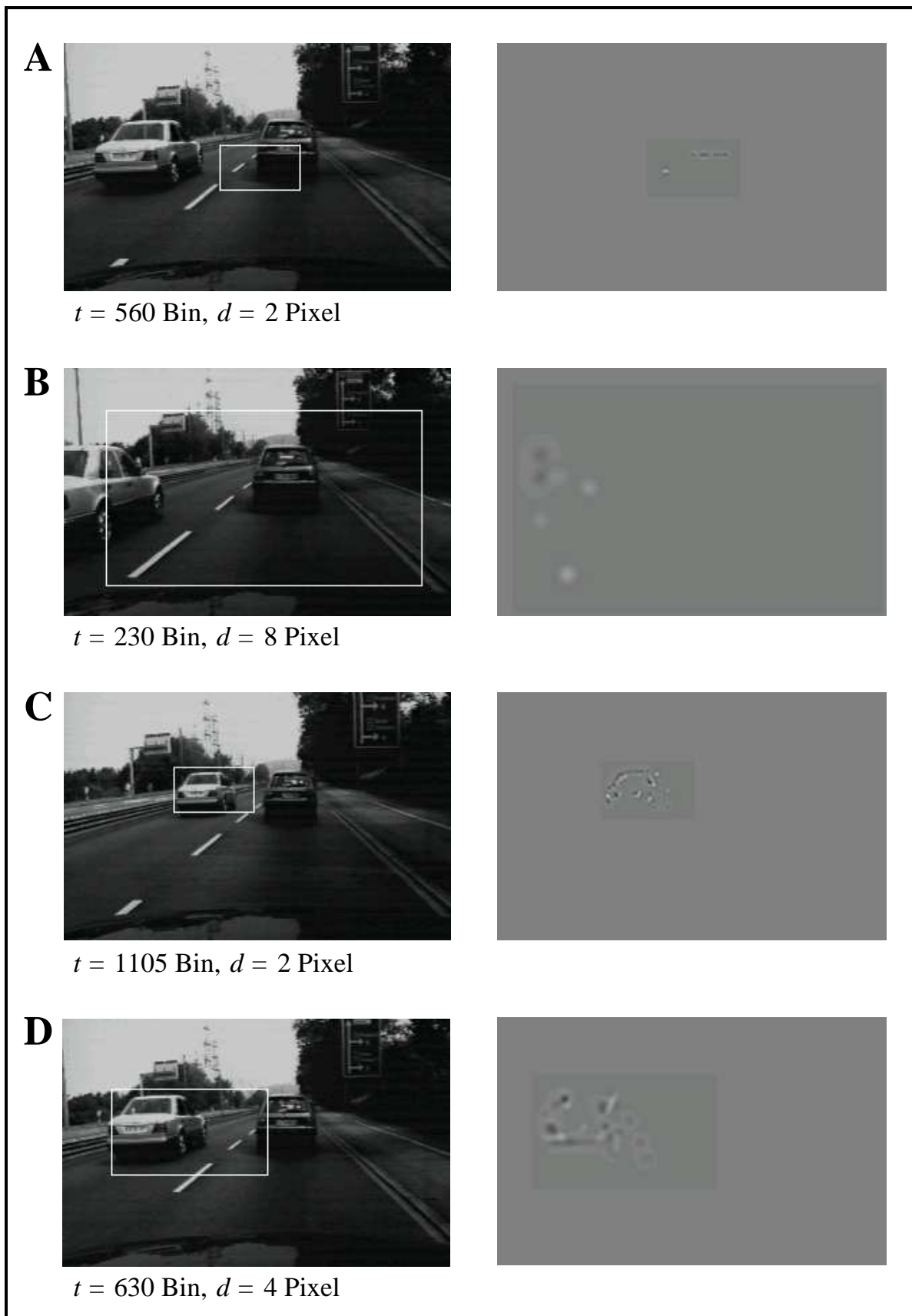


Abb. 3.28: Vorverarbeitung der Sequenz „Gerade“. Dargestellt sind jeweils links das Originalbild mit Abtastbereich (weißer Rahmen) und rechts die rückprojizierten rezeptiven Felder der aktiven Y-Zellen. Die Sequenz „Gerade“ wurde von der FhG-IITB-Karlsruhe zur Verfügung gestellt.

3.2 Bewegungsdetektion mit impulskodierenden neuronalen Netzen

3.2.1 Modell zur Bewegungsdetektion

Das Ziel der Bewegungsverarbeitung ist die Bestimmung der lokalen, zweidimensionalen Bewegungsrichtung sich bewegender Objekte. Das hier vorgestellte Modell (Abb. 3.29) ist unterteilt in ein Vorverarbeitungsmodul und ein Modul zur Extraktion der Bewegungsrichtung. Die Vorverarbeitung findet außerhalb des Acceleratorsystems statt und hat die Aufgabe, eine Datenreduktion durchzuführen. Bei der Datenreduktion werden zwei Fakten ausgenutzt: Da in homogenen Flächen keine lokale Bewegung bestimmt werden kann, beschränkt sich die Bewegungsverarbeitung auf Intensitätskontraste (Abb. 3.29 1). Für die Bewegungsverarbeitung sind nur die sich bewegenden Intensitätskontraste wichtig (Abb. 3.29 2). Stationäre Intensitätskontraste gehen nicht in die Bewegungsverarbeitung ein. Die Beschränkung der Weiterverarbeitung auf sich bewegende Intensitätskontrastkanten anstelle des Grauwertbildes bedeutet eine enorme Datenreduktion, die im Hinblick auf die Merkmalsextraktion und die begrenzte Aufnahmekapazität der Hardware notwendig ist. Die Vorverarbeitung erfolgt in teilweiser Analogie zum biologischen Sehsystem. Die Ausgabe der Vorverarbeitung entspricht in mehreren Eigenschaften prinzipiell der der retinalen Ganglienzellen vom Typ Y.

Das Grundelement des hier vorgestellten Modelles zur Bewegungsverarbeitung ist ein Marburger Modellneuron, das als bilokaler Richtungsdetektor betrieben wird. An jedem Abtastpunkt stehen zwei Sätze Richtungsdetektoren zur Verfügung (Abb. 3.29 3): ein Satz für die Verarbeitung im ON-Pfad und ein zweiter Satz für den OFF-Pfad. Jeder Satz besteht aus sechs Richtungsdetektoren, deren Bewegungsrichtungspräferenzen sich um jeweils 60° unterscheiden. Die lokale Bewegungsrichtung der Intensitätskontrastkante spiegelt sich in der Aktivität der Richtungsdetektoren eines Satzes wider. Da die lokalen Richtungsdetektoren dem Korrespondenzproblem unterliegen, sind in der Regel mehrere Richtungsdetektoren eines Satzes aktiv. Um ein eindeutiges Ergebnis bei der lokalen Bewegungsrichtungsbestimmung zu erhalten, sind in diesem Modell vier zusätzliche neuronale Mechanismen implementiert:

- Laterale Linking-Verbindungen zwischen benachbarten Richtungsdetektoren gleicher Bewegungsrichtungspräferenz, unterstützt durch eine lokale Inhibition über ein Interneuron (Kap. 3.2.1.2, Abb. 3.29 3).
- Verzögerte Voraktivierung des benachbarten Modellneurons über Linking-Kopplung in Richtung der geschätzten Bewegungsrichtung (Kap. 3.2.1.3, Abb. 3.29 4a, b).
- Kopplung der Richtungsdetektoren der beiden ON- und OFF-Pfade (Kap. 3.2.1.4, Abb. 3.29 5).
- Interaktion der Richtungsdetektoren unterschiedlicher Auflösung (Kap. 3.3).

Die angegebene Reihenfolge der Mechanismen ist auch als Rangfolge bezüglich ihrer Effektivität zu verstehen. Insbesondere die ersten beiden Mechanismen führen schon zu guten Ergebnissen (Kapitel 3.2.2). Die beiden anderen Mechanismen können je nach Anzahl an zur Verfügung stehenden Modellneuronen optional implementiert werden. So führt zum Beispiel der vierte Mechanismus zu einer Verdopplung der Modellneuronenanzahl respektive zu einer Halbierung der Abtastpunkte.

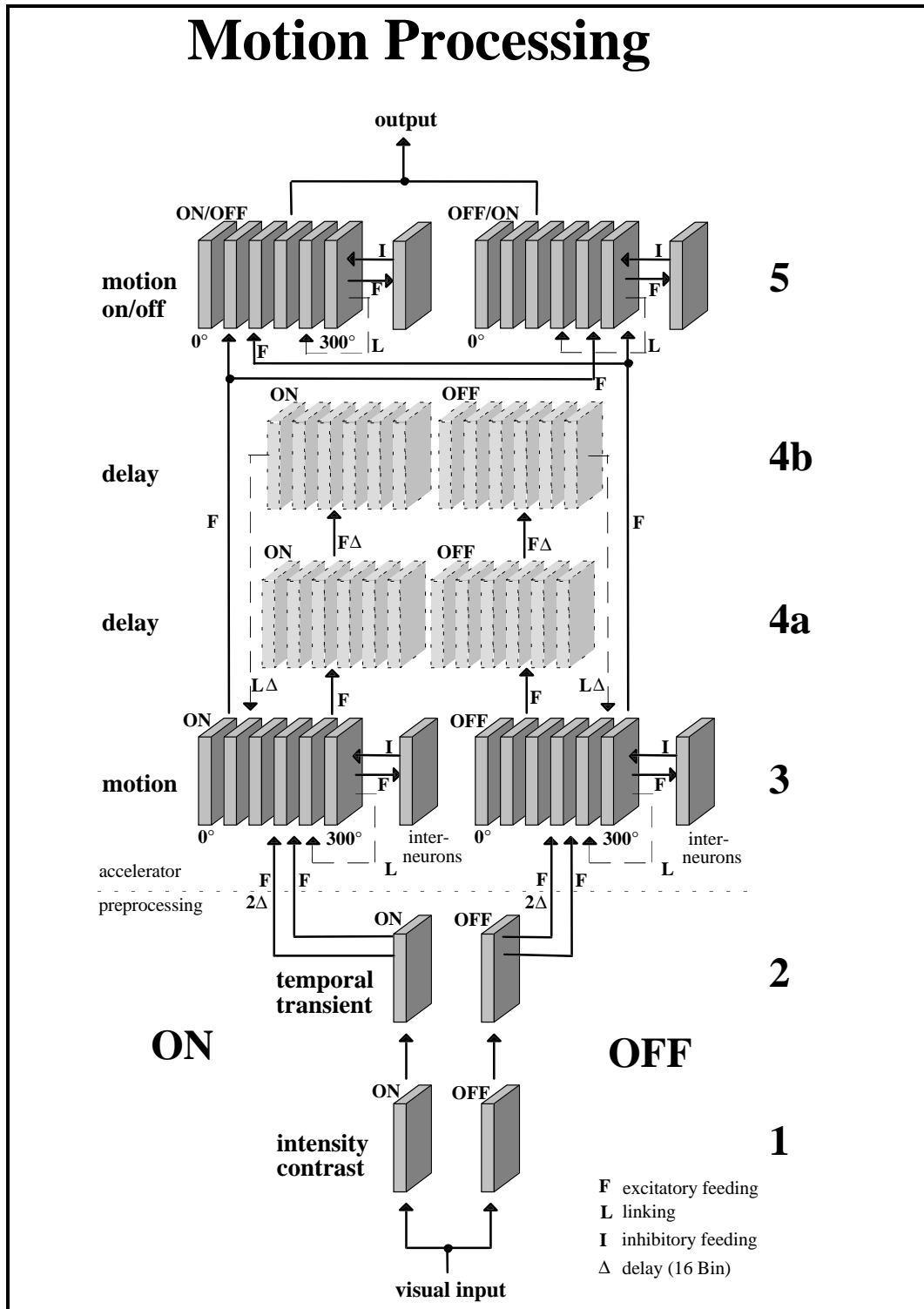


Abb. 3.29: Bewegungsverarbeitung. Die Abbildung zeigt für eine Auflösung die einzelnen Stufen (1 bis 5) der Bewegungsverarbeitung sowie deren Kopplung. Die beiden ersten Stufen befinden sich in der Vorverarbeitung, die nachfolgenden Stufen auf dem Acceleratorsystem. Die Verarbeitung in den ersten vier Stufen erfolgt getrennt in ON- und OFF-Pfad. (1) Intensitätskontrastdetektion, (2) Extraktion zeitlicher Transienten, (3) Bewegungsrichtungsdetektion (sechs Bewegungsrichtungen), (4a, b) Modellneuronen zur verzögerten Voraktivierung der Richtungsdetektoren, (5) Kopplung der ON- und OFF-Richtungsdetektoren. Die Bewegungsrichtungsdetektion wird durch laterale Linking-Kopplungen in den Richtungsdetektorschichten, verzögerte Voraktivierung über Linking-Kopplungen und der Interaktion zwischen ON- und OFF-Pfad unterstützt (Schott et al. 1996).

3.2.1.1 Modell des Richtungsdetektors

Wie bereits oben erwähnt wurde, ist das Grundelement der Bewegungsverarbeitung ein Marburger Modellneuron, das als bilokaler Richtungsdetektor betrieben wird. Zur Ermittlung der lokalen Bewegungsrichtung erhält das Modellneuron seine Eingabe von zwei vorgeschalteten retinalen Ganglienzellen mit benachbarten rezeptiven Feldern über zwei Feeding-Eingänge (Abb. 3.30). Befänden sich verschwindend kleine, nicht überlappende rezeptive Felder in einem Abstand l zueinander, dann könnte bei Verzögerung eines der beiden Eingänge um Δ' die Geschwindigkeit $v = l/\Delta'$ bestimmt werden. Aufgrund der räumlichen Tiefpaßfilterung in der Vorverarbeitung besitzen die rezeptiven Felder eine nicht zu vernachlässigende Größe, wodurch zwar eine exakte Geschwindigkeitsbestimmung nicht mehr möglich ist, was aber auch die Anzahl an Richtungsdetektoren reduziert.

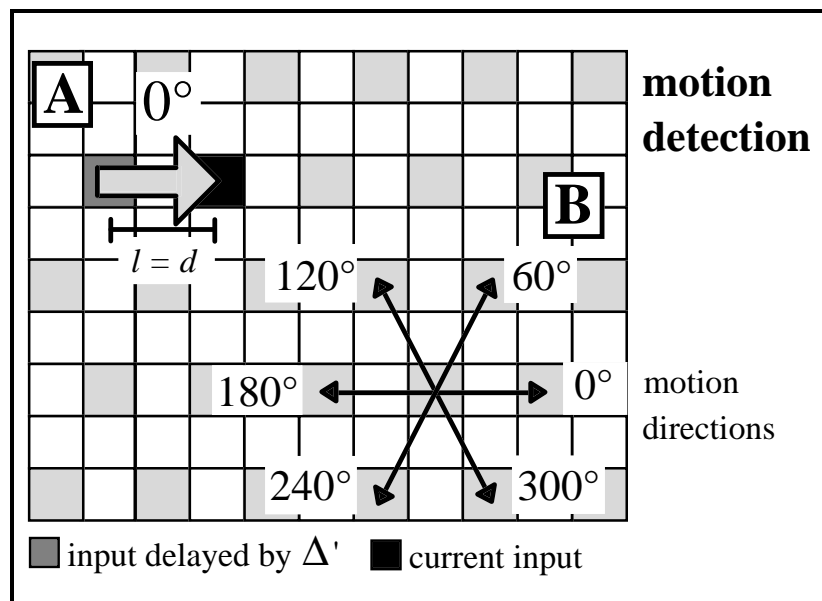
Zur Bewegungsrichtungsbestimmung wurden nur retinale Ganglienzellen mit benachbarten rezeptiven Feldern herangezogen, so daß im hexagonalen Gitter sechs fundamentale Bewegungsrichtungen bestimmt werden können (Gleichungen 3.20 bis 3.26). Das bestimmbare Geschwindigkeitsintervall ist abhängig von der räumlichen Gitterweite. Je gröber die Auflösung ist, um so höher sind die detektierbaren Geschwindigkeiten und umso größer die Geschwindigkeitsintervalle. Die Verzögerung Δ' entspricht dem Zeitintervall T zwischen zwei aufeinanderfolgenden Bildern. Eine Veränderung von Δ' führt natürlich zu einer Verschiebung der auflösungsabhängigen Geschwindigkeitsintervalle, was in technischen Systemen vorteilhaft ausgenutzt werden kann (siehe Diskussion).

Das Gewicht w der Verschaltung zwischen dem Modellneuron Y (i, j) der retinalen Ganglienzellschicht Y wird unverzögert mit dem Modellneuron M (k, l) der Richtungsdetektorschicht M, ϑ über den Feeding-Eingang $F1$ nach Gleichung 3.20 verschaltet.

$$w_{ijkl}^{Y, M, 0^\circ} = \begin{cases} V^{F1}, & \text{falls } ((i = k) \wedge (j = l)) \\ 0, & \text{sonst} \end{cases} \quad (\text{Gl. 3.20})$$

mit $\vartheta \in \{0^\circ, 60^\circ, 120^\circ, 180^\circ, 240^\circ, 300^\circ\}$

Abb. 3.30: Modell zur Bewegungsrichtungsdetektion.
A: Die Bewegungsrichtungsdetektion (hier 0°) erfolgt durch die Verschaltung zweier retinaler Ganglienzellen auf ein nachgeschaltetes Modellneuron (RF-Mittelpunkt bei schwarzer Markierung). Die dunkelgraue Markierung symbolisiert den um Δ' verzögerten und die schwarze Markierung den unverzögerten Eingang.
B: Die Kopplung der retinalen Ganglienzellen mit benachbarten rezeptiven Feldern führt im hexagonalen Abtastgitter zu sechs fundamentalen Bewegungsrichtungen.



i, j, k, l stammen aus den jeweiligen räumlichen Abtastbereichen. Ein Gewicht von $w = 0$ bedeutet hier und in allen weiteren Gleichungen, daß das Gewicht aus Gründen der Rechenzeit im Gewichtespeicher des Accelerators nicht vorhanden ist. Für Gewichte der um Δ verzögerten Eingänge gelten die nachfolgenden Gleichungen 3.21 bis 3.26. Die in allen Simulationen verwendeten Parameter der Richtungsdetektoren sind in Tabelle 3.3 aufgeführt.

$${}_{\text{feedl}}^{Y\Delta} w_{ijkl}^{M,0^\circ} = \begin{cases} {}^M V^{F1}, & \text{falls } ((i = k) \wedge (j = (l - d))) \\ 0, & \text{sonst} \end{cases} \quad (\text{Gl. 3.21})$$

$${}_{\text{feedl}}^{Y\Delta} w_{ijkl}^{M,60^\circ} = \begin{cases} {}^M V^{F1}, & \text{falls } ((i = (k + d)) \wedge (j = (l - d / 2))) \\ 0, & \text{sonst} \end{cases} \quad (\text{Gl. 3.22})$$

$${}_{\text{feedl}}^{Y\Delta} w_{ijkl}^{M,120^\circ} = \begin{cases} {}^M V^{F1}, & \text{falls } ((i = (k + d)) \wedge (j = (l + d / 2))) \\ 0, & \text{sonst} \end{cases} \quad (\text{Gl. 3.23})$$

$${}_{\text{feedl}}^{Y\Delta} w_{ijkl}^{M,180^\circ} = \begin{cases} {}^M V^{F1}, & \text{falls } ((i = k) \wedge (j = (l + d))) \\ 0, & \text{sonst} \end{cases} \quad (\text{Gl. 3.24})$$

$${}_{\text{feedl}}^{Y\Delta} w_{ijkl}^{M,240^\circ} = \begin{cases} {}^M V^{F1}, & \text{falls } ((i = (k - d)) \wedge (j = (l + d / 2))) \\ 0, & \text{sonst} \end{cases} \quad (\text{Gl. 3.25})$$

$${}_{\text{feedl}}^{Y\Delta} w_{ijkl}^{M,300^\circ} = \begin{cases} {}^M V^{F1}, & \text{falls } ((i = (k - d)) \wedge (j = (l - d / 2))) \\ 0, & \text{sonst} \end{cases} \quad (\text{Gl. 3.26})$$

Mit Hinblick auf die von den retinalen Ganglienzellen erzeugten kurzen Aktionspotentialfolgen wird die Feeding-Verstärkung ${}^M V^{F1}$ klein gegenüber dem Schwellenoffset ${}^M \theta_0$ gewählt, so daß bei einer relativ kleinen Feeding-Zeitkonstanten ${}^M \tau^{F1}$ beide Feeding-Eingänge des Modellneurons durch eine Aktionspotentialfolge ausreichend stark gereizt werden müssen, damit der Richtungsdetektor aktiv wird. Sowohl die Schwellenverstärkung ${}^M V^\theta$ des Richtungsdetektors als auch die Schwellenzeitkonstante ${}^M \tau^\theta$ sind klein, damit das Modellneuron eine ausreichende Anzahl an Aktionspotentialen generieren kann. ${}^M V^\theta = 8$ ist die kleinste nicht verschwindende Schwellenverstärkung des Acceleratorsystems. Die Feeding-Zeit-konstante ${}^M \tau^{F1}$ muß klein gewählt werden, damit die Aktivierung des Modellneurons schnell genug abklingt. Ansonsten würde es zu einer räumlichen Verschmierung der Objektkanten kommen.

Feeding-Gewicht Richtungsdetektor ${}^M V^{F1}$	1,75
Feeding-Zeitkonstante ${}^M \tau^{F1}$	5 Bin
Schwellenverstärkung Richtungsdetektor ${}^M V^\theta$	8
Schwellenzeitkonstante ${}^M \tau^\theta$	3 Bin
Schwellenoffset ${}^M \theta_0$	4

Tab. 3.3: Parameter der Richtungsdetektoren. Die Parameter sind sowohl für alle Auflösungen als auch für ON- und OFF-Pfad gleich.

Die Parameter und die Verschaltungen gelten sowohl für den ON- als auch für den OFF-Verarbeitungspfad. Die Verarbeitung wird zunächst getrennt vollzogen, um das Korrespondenzproblem zu minimieren. Diese Vorgehensweise wird durch das in Kap. 2.1.3 dargestellte psychophysische Experiment von Shechter und Hochstein (1990) unterstützt.

Der vorgestellte Richtungsdetektor unterscheidet sich vom Reichardt-Detektor in zwei wesentlichen Punkten: Zum einen erfolgt die Verrechnung der beiden Eingänge nicht multiplikativ, sondern additiv, zum anderen werden stehende Intensitätskontrastkanten bereits vor der Richtungsverarbeitung eliminiert. Der erste Unterschied bringt den Vorteil mit sich, daß der mögliche Membranpotentialbereich bei einer additiven Verschaltung deutlich kleiner als bei einer multiplikativen Verschaltung ist. Als Beispiel hierfür sei auf die Abhängigkeit der Aktivität vom Intensitätskontrast hingewiesen (Kap. 3.1.2). Ferner besitzt das verwendete Acceleratorsystem nur einen Linking-Eingang, wodurch es nicht möglich ist, gleichzeitig eine Multiplikation und eine Modulation zu verwirklichen. Da die modulatorischen Eigenschaften der Linking-Verbindungen bei der Bewegungsanalyse entscheidend sind, muß die Verschaltung des Richtungsdetektors additiv erfolgen. Der zweite Unterschied impliziert natürlich eine enorme Aktivitätsreduzierung und eine Verringerung des Korrespondenzproblems.

3.2.1.2 Laterale, unverzögerte Kopplung der Richtungsdetektoren

Wie in Kapitel 2.1.1 dargestellt wurde, unterliegen bilokale Richtungsdetektoren dem Korrespondenzproblem. Bereits eine einzelne Kante führt zu einem mehrdeutigen Ergebnis bei der Bewegungsrichtungsanalyse. Die menschliche Wahrnehmung der Bewegungsrichtung ist aber eindeutig. Die im folgenden beschriebenen lateralen Kopplungen sollen die erste Stufe zu einem solchen eindeutigen Richtungsdetektionsergebnis liefern.

Die lateralen Kopplungen erfolgen über Linking-Verbindungen und werden durch *die Gestaltprinzipien des guten Verlaufes, der Nähe und des gemeinsamen Schicksals* motiviert (siehe Kap. 2.1.2.1 und 2.1.2.3). Die Kopplungen finden über Linking-Verbindungen statt, damit die rezeptive Feldstruktur nicht zerstört wird. Außerdem muß die Kopplung im Hinblick auf die Bewegung schnell erfolgen. Da Linking-Verbindungen modulatorisch und nicht additiv wirken, kommt es zu einer schnellen, kurzzeitigen Synchronisation neuronaler Aktivität.

Das Gestaltprinzip des guten Verlaufes spiegelt sich in den Verbindungen entlang einer Kontrastkante wider. Diese Verbindungen sind aber zum einen räumlich begrenzt und zum anderen auf Richtungsdetektoren der gleichen Richtungspräferenz beschränkt, was in den Gestaltprinzipien der Nähe bzw. des gleichen Schicksals begründet ist. Die Abbildung 3.31 zeigt die lateralen Linking-Verschaltungen anhand eines 0°-Richtungsdetektors. Es sind unverzögerte und verzögerte (Kapitel 3.2.1.3) Kopplungen implementiert. Der 0°-Richtungsdetektor ist mit den nächsten 0°-Richtungsdetektoren in Richtung 60°, 120°, 240° und 300° unverzögert gekoppelt. Hierbei wird von einem möglichen Kantenverlauf in diese Bereiche ausgegangen. Eine Kopplung in Richtung 180° erfolgt nicht, da dort keine transiente Antwort einer Kante auftreten kann. Wird eine 0° oder 180° orientierte Kante in Richtung 0° bewegt, dann verändern sich nur die Enden der Kanten. Der gesamte innere Teil ist nicht im Transientensystem repräsentiert. Eine Kopplung in die der Bewegung entgegengesetzten

Richtung hätte vielmehr einen unerwünschten „Aktivitätsschweif“ zur Folge, so daß eine Inhibition in diese Richtung vorteilhaft wäre. Diese Inhibition erfolgt über das weiter unten beschriebene Interneuron.

Für die lateralen, unverzögerten Kopplungen zwischen Modellneuron $M(i, j)$ der Richtungspräferenz ϑ' und dem Modellneuron $M(k, l)$ der Richtungspräferenz ϑ ergeben sich die Linking-Verbindungen w nach den Gleichungen 3.27 bis 3.29.

$${}_{\text{link}} w_{ijkl}^{M, \vartheta'=\vartheta} = \begin{cases} {}^M V^L, & \text{falls } ((i = (k + d)) \wedge (j = (l \pm d/2))) \vee \\ & ((i = (k - d)) \wedge (j = (l \pm d/2))) \\ 0, & \text{sonst} \end{cases} \quad (\text{Gl. 3.27})$$

für $\vartheta \in \{0^\circ, 180^\circ\}$

(Gl. 3.28)

$${}_{\text{link}} w_{ijkl}^{M, \vartheta'=\vartheta} = \begin{cases} {}^M V^L, & \text{falls } ((i = k) \wedge (j = (l \pm d))) \vee ((i = (k + d)) \wedge (j = (l + d/2))) \vee \\ & ((i = (k - d)) \wedge (j = (l - d/2))) \\ 0, & \text{sonst} \end{cases}$$

für $\vartheta \in \{60^\circ, 240^\circ\}$

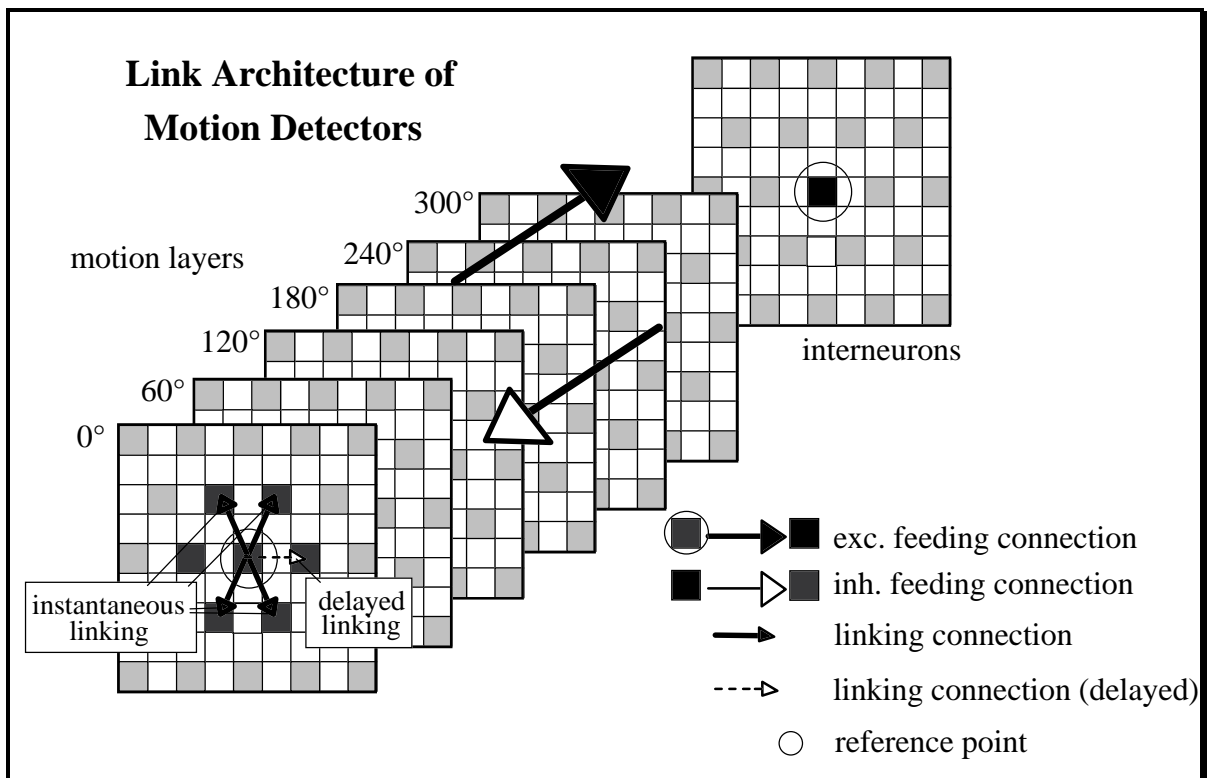


Abb. 3.31: Link-Architektur der Richtungsdetektoren. Die Abbildung zeigt die sechs Richtungsdetektorschichten und eine Schicht Interneuronen. Ferner ist für einen 0°-Richtungsdetektor die laterale Verschaltung über Linking-Verbindungen dargestellt. Aus Gründen der Aktivitätsbeschränkung existiert eine lokale Inhibition über ein Interneuron. Nähere Erläuterungen im Text.

(Gl. 3.29)

$$M_{\text{link}}^{\vartheta=\vartheta} w_{ijkl}^{\vartheta} = \begin{cases} V^L, & \text{falls } ((i = k) \wedge (j = (l \pm d))) \vee ((i = (k + d)) \wedge (j = (l - d/2))) \vee \\ & ((i = (k - d)) \wedge (j = (l + d/2))) \\ 0, & \text{sonst} \end{cases}$$

für $\vartheta \in \{120^\circ, 300^\circ\}$

i, j, k und l sind wieder aus den jeweiligen Abtastbereichen zu wählen. d gibt die räumliche Abtastweite an.

Diese lateralen Kopplungen erscheinen auch biologisch plausibel. Da Modellneuronen mit benachbarten rezeptiven Feldern bei Aktivierung durch dieselbe Kontrastkante nahezu koinzident aktiv werden, sollten Verbindungen zwischen ihnen nach der Hebb'schen Lernregel (Hebb 1949) verstärkt werden. Die Kopplungen wären unter Beachtung der Aktivität zwischen Modellneuronen gleicher Richtungspräferenz am stärksten. Eine Erhöhung der Verbindungsstärke zwischen Modellneuronen mit entgegengesetzten Richtungspräferenzen wäre am unwahrscheinlichsten.

Durch die lateralen Linking-Verbindungen erhöht sich die Aktivität der Modellneuronen, deren benachbarte Modellneuronen ebenfalls von einer sich bewegenden Kante mit der gleichen Bewegungsrichtung stimuliert werden. Richtungsdetektoren, denen solche benachbarten Richtungsdetektoren fehlen, aber aufgrund des Korrespondenzproblems trotzdem aktiv sind, erhöhen ihre Aktivität hingegen nicht. Ein eindeutiges Ergebnis bei der lokalen Bewegungsanalyse wird erreicht, indem die Membranpotentiale der schwach aktiven Richtungsdetektoren über eine zusätzliche aktivitätsabhängige Inhibition unter ihren Schwellenoffset gedrückt werden. Dies entspricht einem dynamischen *winner-take-all*-Mechanismus. Die Inhibition erfolgt über Interneuronen einer zusätzlichen Schicht M, Inter, die ihre Eingabe von Richtungsdetektoren mit gleichem RF-Mittelpunkt erhalten (Gl. 3.30). Das Interneuron wirkt inhibitorisch auf diese und alle nächsten Nachbarn zurück (Gl. 3.31). Die Verschaltung zwischen Richtungsdetektoren und Interneuronen ist ebenfalls in Abbildung 3.31 dargestellt.

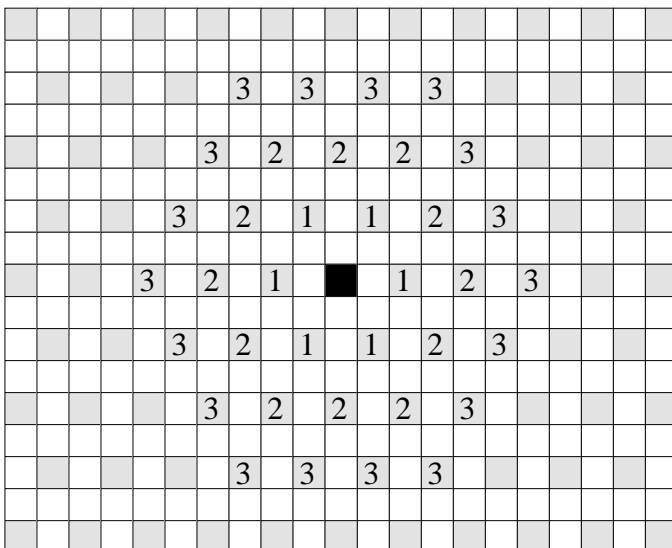


Abb. 3.32: Nachbarschaften im hexagonalen Gitter. Dargestellt sind die um das Zentrum (schwarze Markierung) befindlichen drei nächsten Hexagone (Bezeichnungen 1 bis 3) in der Auflösung mit Abtastweite $d = 2$ Pixel. Zu „hexagon 2“ gehören beispielsweise alle Modellneuronen, deren zugehörige Gitterplätze mit „2“ gekennzeichnet sind. Das nächste „hexagon $n+1$ “ wird durch alle direkt an „hexagon n “ angrenzende freie Gitterplätze gebildet.

Linking-Gewicht Richtungsdetektor ${}^M V^L$	1,0
Linking-Zeitkonstante ${}^M \tau^L$	5 Bin
Inhibition-Gewicht Richtungsdetektor ${}^M V^I$	5,0
Inhibition-Zeitkonstante ${}^M \tau^I$	7,5 Bin
Feeding-Gewicht Interneuron ${}^{M, Inter} V^{F1}$	2,5
Feeding-Zeitkonstante ${}^{M, Inter} \tau^{F1}$	5 Bin
Schwellenverstärkung Interneuron ${}^{M, Inter} V^\theta$	8
Schwellenzeitkonstante ${}^{M, Inter} \tau^\theta$	3 Bin
Schwellenoffset ${}^{M, Inter} \theta_0$	4

Tab. 3.4: Parameter der Richtungsdetektoren und Interneuronen. Die Parameter sind sowohl für alle Auflösungen als auch für ON- und OFF-Pfad gleich.

$${}^{M, \vartheta'}_{feed1} w_{ijkl}^{M, Inter} = \begin{cases} {}^{M, Inter} V^{F1}, & \text{falls } ((i == k) \wedge (j == l)) \\ 0, & \text{sonst} \end{cases} \quad (\text{Gl. 3.30})$$

$$\text{für } \vartheta' \in \{0^\circ, 60^\circ, 120^\circ, 180^\circ, 240^\circ, 300^\circ\}$$

$${}^{M, Inter}_{inhib} w_{ijkl}^{M, \vartheta} = \begin{cases} {}^M V^I, & \text{falls } ((i == k) \wedge (j == l)) \vee \text{hexagon}_d 1 \\ 0, & \text{sonst} \end{cases} \quad (\text{Gl. 3.31})$$

$$\text{für } \vartheta \in \{0^\circ, 60^\circ, 120^\circ, 180^\circ, 240^\circ, 300^\circ\}$$

„ $\text{hexagon}_d 1$ “ bezeichnet hierbei die nächsten Nachbarn im hexagonalen Gitter der Abtastweite d . Die Abbildung 3.32 zeigt am Beispiel der Abtastweite $d = 2$ Pixel die in Hexagonen zusammengefaßten Nachbarschaften „ $\text{hexagon}_2 1$ “, „ $\text{hexagon}_2 2$ “ und „ $\text{hexagon}_2 3$ “. Die Parameter des Interneurons und die Parameter des inhibitorischen Dendriten sowie des Linking-Eingangs der Richtungsdetektoren sind in Tabelle 3.4 aufgelistet.

Die Inhibition könnte auch direkt erfolgen, was aber einen erhöhten Verschaltungsaufwand zur Folge hätte. Die inhibitorische Verschaltung eines Richtungsdetektorsatzes erfordert sechs exzitatorische Verbindungen zu dem Interneuron und 42 inhibitorische Verbindungen zu den Richtungsdetektoren. Ohne Interneuron sind 252 Verbindungen nötig. Zusätzlich können die Eigenschaften des Interneurons ausgenutzt werden. Eine zu schwache Aktivität der Richtungsdetektoren führt zum Beispiel nicht zu einer Aktivierung des Interneurons und somit auch zu keiner Inhibition und einer weiteren Aktivitätsabschwächung.

3.2.1.3 Laterale, verzögerte Kopplung der Richtungsdetektoren

Bewegt sich eine Kante in eine bestimmte Bewegungsrichtung, so kann in der Regel davon ausgegangen werden, daß sich die Bewegungsrichtung – im weiteren Sinne das „gemeinsame Schicksal“ – innerhalb weniger Millisekunden nicht oder nur wenig ändert. Unter Verwendung der von den Modellneuronen gemachten Schätzung der Bewegungsrichtung ist eine Voraktivierung benachbarter Modellneuronen gleicher Richtungspräferenz sinnvoll. Da die Kante bis zum Erreichen der nächsten RF-Mittelpunkte eine bestimmte Zeit braucht, sollte die Voraktivierung verzögert erfolgen. Durch diese Verzögerung kommt es neben der räumlichen auch zu einer zeitlichen Integration der Bewegungsinformation. Je länger die Verzögerung ist, um so mehr kann zeitlich integriert werden. Bei ausreichend großer Verzögerung könnte die detektierte Bewegungsrichtung eines Objektes auch bei kurzzeitiger Objektverdeckung erhalten bleiben.

In technischen Anwendungen sind aber aus Speicherplatzgründen nur kurze Verzögerungen möglich, wodurch auch nur kurzreichweitige Verbindungen möglich sind. Der Paderborner Accelerator bietet die Möglichkeit, die Aktionspotentialausgabe, einer axonalen Verzögerung entsprechend, um 16 Bin zu verzögern. Eine Verzögerung um ein Intervall zwischen zwei aufeinanderfolgenden Bildern benötigt damit bereits zwei Modellneuronen. Die Verschaltung wurde bereits in den Abbildungen 3.29 und 3.31 dargestellt. Die Schichten 4a und 4b des Modells aus Abbildung 3.29 dienen lediglich der Verzögerung der Ausgabe der Modellneuronen aus Schicht 3. Die Voraktivierung beschränkt sich aus den oben genannten Gründen auf das Modellneuron, dessen rezeptives Feld wahrscheinlich als nächstes von der Kante durchlaufen wird und dessen Richtungspräferenz der geschätzten Bewegungsrichtung entspricht.

Der Richtungsdetektor $M(i, j)$ mit der Richtungspräferenz ϑ' wird mit dem Verzögerungsneuron $MDa(k, l)$ der Schicht MDa , ϑ über eine Feeding-Verbindung nach Gleichung 3.32 gekoppelt. Entsprechend werden Verzögerungsneuronen der Schicht MDa , ϑ' mit Verzögerungsneuronen der Schicht MDb , ϑ verbunden (Gl. 3.33). Die Modellneuronen der beiden Verzögerungsschichten geben die Aktionspotentiale mit jeweils 16 Bin Verzögerung weiter. Die Voraktivierung der Richtungsdetektoren erfolgt dann in Abhängigkeit von der Richtungspräferenz um das Zeitintervall Δ zwischen zwei aufeinanderfolgenden Bildern nach den Gleichungen 3.34 bis 3.39.

$${}_{\text{feed1}}^{M, \vartheta'=\vartheta} w_{ijkl}^{MDa, \vartheta} = \begin{cases} {}^{M, \text{delay}} V^{F1}, & \text{falls } ((i == k) \wedge (j == l)) \\ 0, & \text{sonst} \end{cases} \quad (\text{Gl. 3.32})$$

$${}_{\text{feed1}}^{MDa, \vartheta'=\vartheta} w_{ijkl}^{MDb, \vartheta} = \begin{cases} {}^{M, \text{delay}} V^{F1}, & \text{falls } ((i == k) \wedge (j == l)) \\ 0, & \text{sonst} \end{cases} \quad (\text{Gl. 3.33})$$

für die Gleichungen 3.32 und 3.33 ist $\vartheta \in \{0^\circ, 60^\circ, 120^\circ, 180^\circ, 240^\circ, 300^\circ\}$.

$${}_{\text{link}}^{MDb, 0^\circ} w_{ijkl}^{M, 0^\circ} = \begin{cases} {}^{M, \text{vorak}} V^L, & \text{falls } ((i == k) \wedge (j == (l - d))) \\ 0, & \text{sonst} \end{cases} \quad (\text{Gl. 3.34})$$

$${}_{\text{link}}^{M\text{Db},60^\circ} w_{ijkl}^{M,60^\circ} = \begin{cases} {}^{\text{M, vorak}} V^L, & \text{falls } ((i == (k + d)) \wedge (j == (l - d/2))) \\ 0, & \text{sonst} \end{cases} \quad (\text{Gl. 3.35})$$

$${}_{\text{link}}^{M\text{Db},120^\circ} w_{ijkl}^{M,120^\circ} = \begin{cases} {}^{\text{M, vorak}} V^L, & \text{falls } ((i == (k + d)) \wedge (j == (l + d/2))) \\ 0, & \text{sonst} \end{cases} \quad (\text{Gl. 3.36})$$

$${}_{\text{link}}^{M\text{Db},180^\circ} w_{ijkl}^{M,180^\circ} = \begin{cases} {}^{\text{M, vorak}} V^L, & \text{falls } ((i == k) \wedge (j == (l + d))) \\ 0, & \text{sonst} \end{cases} \quad (\text{Gl. 3.37})$$

$${}_{\text{link}}^{M\text{Db},240^\circ} w_{ijkl}^{M,240^\circ} = \begin{cases} {}^{\text{M, vorak}} V^L, & \text{falls } ((i == (k - d)) \wedge (j == (l + d/2))) \\ 0, & \text{sonst} \end{cases} \quad (\text{Gl. 3.38})$$

$${}_{\text{link}}^{M\text{Db},300^\circ} w_{ijkl}^{M,300^\circ} = \begin{cases} {}^{\text{M, vorak}} V^L, & \text{falls } ((i == (k - d)) \wedge (j == (l - d/2))) \\ 0, & \text{sonst} \end{cases} \quad (\text{Gl. 3.39})$$

Die Parameter der Verzögerungsneuronen und der Voraktivierung sind in Tabelle 3.5 aufgeführt. Das Linking-Gewicht der Voraktivierung wird größer gewählt als das der unverzögerten Verbindung, da die Voraktivierung in der Regel nur von einem Modellneuron ausgeht, wohingegen mehrere aktive Nachbarneuronen bei der unverzögerten, lateralen Kopplung vorhanden sind. Die Verzögerungsneuronen sind technisch bedingt und haben nur die Aufgabe, die Aktionspotentiale verzögert weiterzuleiten. Sie haben bezüglich der visuellen Informationsverarbeitung keine weitere funktionelle Relevanz.

Das voraktivierte Modellneuron eines Richtungsdetektorsatzes hat gegenüber den anderen Modellneuronen ein höheres Membranpotential. Durch die implementierte lokale Inhibition setzt sich das voraktivierte Modellneuron gegenüber den anderen durch.

Eine unverzögerte Voraktivierung wäre nach der Hebbschen Lernregel (Hebb 1949) nicht zu erwarten, da die zugehörigen Modellneuronen nicht gleichzeitig aktiv sind. Stünden zahlreiche Verbindungen mit unterschiedlichen Verzögerungen zwischen diesen Modellneuronen zur Verfügung, so würden nach Hebbschem Lernen nur die Verbindungen mit einer Verzögerung um Δ fortbestehen.

Linking-Gewicht Richtungsdetektor ${}^{\text{M, vorak}} V^L$	2,0
Linking-Zeitkonstante ${}^{\text{M}} \tau^L$	5 Bin
Feeding-Gewicht Verzögerungsneuron ${}^{\text{M, delay}} V^{\text{F1}}$	7,5
Feeding-Zeitkonstante ${}^{\text{M, delay}} \tau^{\text{F1}}$	5 Bin
Schwellenverstärkung Verzögerungsneuron ${}^{\text{M, delay}} V^\theta$	8
Schwellenzeitkonstante ${}^{\text{M, delay}} \tau^\theta$	5 Bin
Schwellenoffset ${}^{\text{M, delay}} \theta_0$	4

Tab. 3.5: Parameter der Voraktivierung und der Verzögerungsneuronen. Die Parameter sind sowohl für alle Auflösungen als auch für ON- und OFF-Pfad gleich.

3.2.1.4 Kopplung von ON- und OFF-Pfad

Die Bewegung einer Intensitätskontrastkante führt zu einer Aktivierung der retinalen Ganglienzellen sowohl im ON- als auch im OFF-Verarbeitungspfad. Abhängig von der Kantenpolarität zur Bewegungsrichtung kommt es in einem der beiden Pfade zu einer starken und in dem anderen zu einer schwachen Aktivierung (Kap. 3.1.2). Die Richtungsdetektoren des ON- und des OFF-Pfades verhalten sich ebenso (Abb. 3.33), so daß eine Kopplung der beiden Verarbeitungspfade das System störunanfälliger macht und zu einem eindeutigen Bewegungsanalyseergebnis, insbesondere bei der Verarbeitung realer Sequenzen, beitragen kann.

Die Richtungsdetektoren werden paarweise, koaxial in Richtungspräferenz zu einem neuen Richtungsdetektor zusammengefaßt. Für eine Bewegungsrichtung sind zwei neue Richtungsdetektoren nötig: ein ONOFF- und ein OFFON-Richtungsdetektor. Die Verschaltung der ONOFF-Richtungsdetektoren ergibt sich aus den Gleichungen 3.40 bis 3.46. Die Parameter sind in Tabelle 3.6 aufgeführt. Die ONOFF-Richtungsdetektoren sind ebenso wie die Richtungsdetektoren lateral über Linking-Verbindungen und über inhibitorisch wirkende Interneuronen gekoppelt. Die Parameter der Interneuronen entsprechen denen aus Tabelle 3.4. Die OFFON-Richtungsdetektoren werden analog konstruiert.

$${}_{\text{feed1}}^{MON, \vartheta' = \vartheta} w_{ijkl}^{MONOFF, \vartheta} = \begin{cases} {}^M, \text{ONOFF} V^{F1}, & \text{falls } ((i == k) \wedge (j == l)) \\ 0, & \text{sonst} \end{cases} \quad (\text{Gl. 3.40})$$

für $\vartheta \in \{0^\circ, 60^\circ, 120^\circ, 180^\circ, 240^\circ, 300^\circ\}$

$${}_{\text{feed1}}^{MOFF, 0^\circ} w_{ijkl}^{MONOFF, 0^\circ} = \begin{cases} {}^M, \text{ONOFF} V^{F1}, & \text{falls } ((i == k) \wedge (j == (l + d))) \\ 0, & \text{sonst} \end{cases} \quad (\text{Gl. 3.41})$$

$${}_{\text{feed1}}^{MOFF, 60^\circ} w_{ijkl}^{MONOFF, 60^\circ} = \begin{cases} {}^M, \text{ONOFF} V^{F1}, & \text{falls } ((i == (k - d)) \wedge (j == (l + d/2))) \\ 0, & \text{sonst} \end{cases} \quad (\text{Gl. 3.42})$$

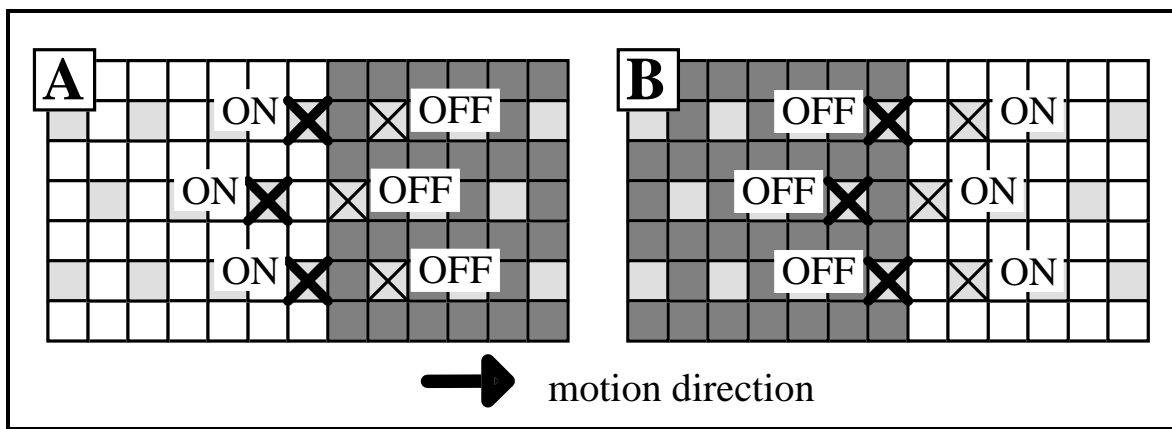


Abb. 3.33: Aktivität der ON- und OFF-Richtungsdetektoren. Die Abbildung zeigt in **A** eine hell/dunkel-Kante und in **B** eine dunkel/hell-Kante, die sich beide in Richtung 0° bewegen und die daraus resultierende Aktivität der ON- und OFF-Richtungsdetektoren. **A** zeigt eine starke Aktivität der ON-Richtungsdetektoren (dicke Kreuze) und eine schwache Aktivierung der OFF-Richtungsdetektoren (dünne Kreuze). Für eine umgekehrte Kantenpolarität ergibt sich, wie in **B** dargestellt, eine Vertauschung der Aktivität zwischen ON- und OFF-Detektoren.

$${}^{\text{MOFF},120^\circ}_{\text{feed1}} w_{ijkl}^{\text{MONOFF},120^\circ} = \begin{cases} {}^{\text{M, ONOFF}} V^{\text{F1}}, & \text{falls } ((i == (k - d)) \wedge (j == (l - d/2))) \\ 0, & \text{sonst} \end{cases} \quad (\text{Gl. 3.43})$$

$${}^{\text{MOFF},180^\circ}_{\text{feed1}} w_{ijkl}^{\text{MONOFF},180^\circ} = \begin{cases} {}^{\text{M, ONOFF}} V^{\text{F1}}, & \text{falls } ((i == k) \wedge (j == (l - d))) \\ 0, & \text{sonst} \end{cases} \quad (\text{Gl. 3.44})$$

$${}^{\text{MOFF},240^\circ}_{\text{feed1}} w_{ijkl}^{\text{MONOFF},240^\circ} = \begin{cases} {}^{\text{M, ONOFF}} V^{\text{F1}}, & \text{falls } ((i == (k + d)) \wedge (j == (l - d/2))) \\ 0, & \text{sonst} \end{cases} \quad (\text{Gl. 3.45})$$

$${}^{\text{MOFF},300^\circ}_{\text{feed1}} w_{ijkl}^{\text{MONOFF},300^\circ} = \begin{cases} {}^{\text{M, ONOFF}} V^{\text{F1}}, & \text{falls } ((i == (k + d)) \wedge (j == (l + d/2))) \\ 0, & \text{sonst} \end{cases} \quad (\text{Gl. 3.46})$$

Linking-Gewicht ONOFF-Richtungsdetektor ${}^{\text{M, ONOFF}} V^{\text{L}}$	1,0
Linking-Zeitkonstante ${}^{\text{M, ONOFF}} \tau^{\text{L}}$	5 Bin
Feeding-Gewicht ${}^{\text{M, ONOFF}} V^{\text{F1}}$	2,0
Feeding-Zeitkonstante ${}^{\text{M, ONOFF}} \tau^{\text{F1}}$	5 Bin
Schwellenverstärkung ONOFF-Richtungsdetektor ${}^{\text{M, ONOFF}} V^{\theta}$	8
Schwellenzeitkonstante ${}^{\text{M, ONOFF}} \tau^{\theta}$	3 Bin
Schwellenoffset ${}^{\text{M, ONOFF}} \theta_0$	4

Tab. 3.6: Parameter der ONOFF-Richtungsdetektoren.

3.2.2 Simulationsergebnisse der Bewegungsdetektion

3.2.2.1 Simulationsergebnisse mit künstlichen Szenen

In diesem Kapitel werden die Eigenschaften der Richtungsdetektoren mittels künstlicher Szenen untersucht. Es wird gezeigt, daß das in 3.2.1 beschriebene Modell zur Bewegungsrichtungsdetektion eine mögliche Verschaltung zur Erklärung der grundlegenden psychophysischen Experimente, dargestellt in Kapitel 2.1, bietet. Insbesondere die Linking-Verbindungen zwischen den Richtungsdetektoren spielen hierbei eine entscheidende Rolle.

Trennung der Bewegungsverarbeitung in ON- und OFF-Pfad

Das in Kapitel 2.1.2.1 beschriebene psychophysische Experiment von Shechter und Hochstein (1990) zeigt, daß Korrespondenz zwischen Flecken in aufeinanderfolgenden Bildern durch eine zunächst getrennte Verarbeitung in einem ON- und einem OFF-Pfad hergestellt werden kann. Diesem Experiment entsprechend werden zwei kleine Flecken (3 Pixel x 3 Pixel) dem Netzwerk in zwei aufeinanderfolgenden Bildern so präsentiert, daß bei gleichem Grauwert der Flecken keine eindeutige Bewegungsrichtung bestimmt werden kann (Abbildung 3.34 A/B), wohingegen eine unterschiedliche Polarität der Flecken (Abbildung 3.34 C/D und E/F) zu einem eindeutigen Bewegungsanalyseergebnis führt.

In Abbildung 3.34 A springen die Flecken ($g_1=50$, $g_2=50$ und Hintergrund $g_H=100$) von den Positionen 1 und 4 auf die Positionen 2 und 3. Eine eindeutige Bewegungsrichtungsbestimmung ist aufgrund fehlender Korrespondenz nicht möglich. So sind Modellneuronen mit RF-Mittelpunkt an Position 2 und einer Richtungspräferenz von 0° und 120° sowie Modellneuronen mit RF-Mittelpunkt an Position 3 und einer Richtungspräferenz von 180° und 300° aktiv (Abb. 3.34 B). Es werden nur Richtungsdetektoren des OFF-Pfades aktiviert.

Haben die beiden Flecken eine unterschiedliche Polarität ($g_1=50$, $g_2=150$ und Hintergrund $g_H=100$), dann wird, wie in der Abbildung 3.34 D und F verdeutlicht, die Bewegungsdetektion in ON- und OFF-Pfad aufgeteilt, wodurch eine eindeutige Bewegungsrichtung bestimmt werden kann. In Abbildung 3.34 C springt der helle Fleck von Position 4 nach 3, was nur zu einer Aktivierung des ON-Modellneurons mit Richtungspräferenz von 180° führt. Das OFF-Modellneuron mit RF-Mittelpunkt an Position 2 zeigt durch seine Aktivierung einen Wechsel des dunklen Fleckes von Position 1 nach 2 an. In Abbildung 3.34 F sind die Modellneuronenaktivitäten für einen Positionswechsel des hellen Fleckes von 4 nach 2 und für einen Positionswechsel des dunklen Fleckes von 1 nach 3 dargestellt. Auch bei dieser Konstellation werden die Bewegungsrichtungen richtig detektiert.

Das neuronale Netz bestand nur aus Richtungsdetektoren (ON- und OFF-Pfad, 6 Richtungen) ohne laterale Kopplungen. Es wurden die Parameter der Tabelle 3.3 benutzt.

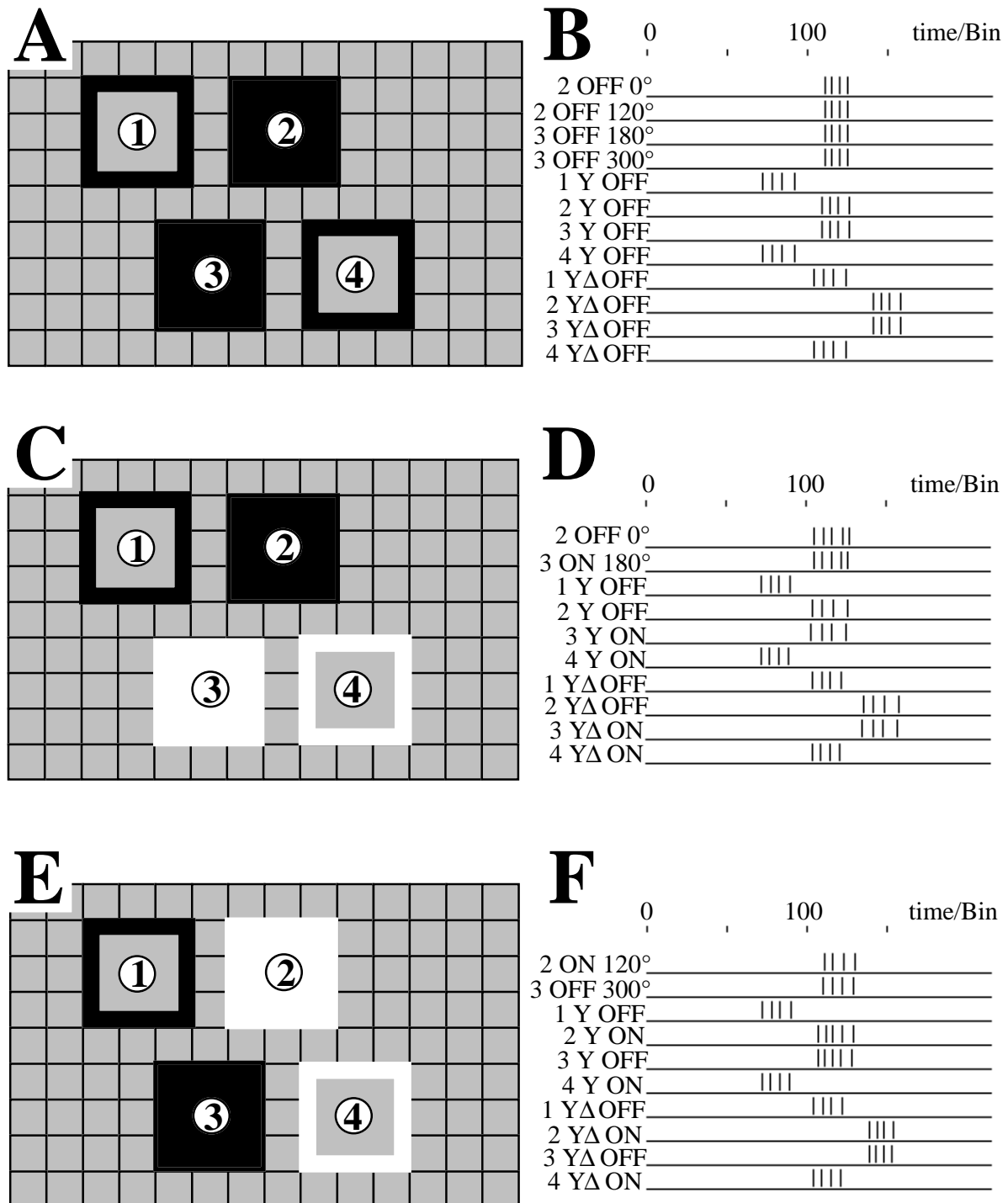


Abb. 3.34: Bewegungsdetektion in ON- und OFF-Pfad. Die Abbildung zeigt für drei verschiedene Fälle den Positionswechsel zweier Flecken von 1 und 4 (leere Rechtecke) nach 2 und 3 (ausgefüllte Rechtecke) und die dadurch verursachten neuronalen Aktivitäten (Richtungsdetektoren des ON- und OFF-Pfades mit angegebener Richtungspräferenz sowie Y-Zellen der Vorverarbeitung; Positionen der RF-Mittelpunkte sind den Bezeichnungen vorangestellt). **A/B:** Aufgrund fehlender Korrespondenz ist die Bewegungsrichtungsdetektion mehrdeutig. **C/D, E/F:** Korrespondenz wird durch unterschiedliche Polarität der Flecken hergestellt, was zu einer eindeutigen Bewegungsdetektion führt.

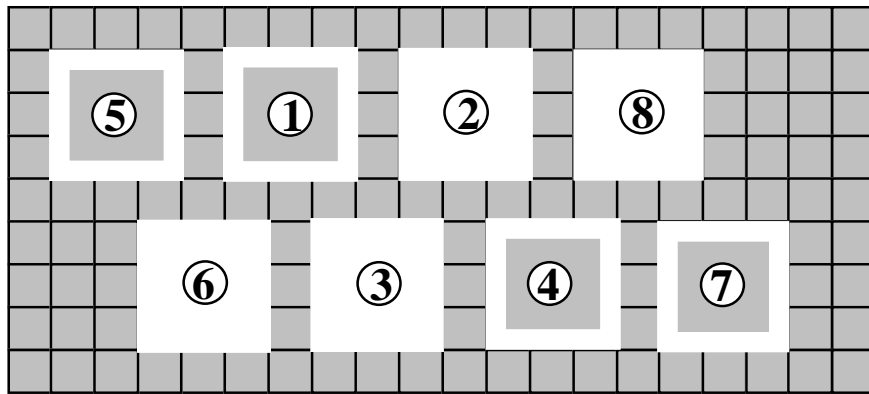


Abb. 3.35: Stimulus zur Untersuchung lateraler Interaktionen zwischen Richtungsdetektoren. In zwei aufeinanderfolgenden Bildern springen die Flecken von den Positionen 5, 1, 4 und 7 (leere Vierecke) zu den Positionen 6, 3, 2 und 8. In der Simulation wird die Anzahl der Flecken auf der linken Seite (Bewegung nach unten) und auf der rechten Seite (Bewegung nach oben) variiert. Der Reiz geht auf psychophysische Untersuchungen zur räumlichen Integration der Bewegungsinformation von Yuille und Grzywacz (1988) zurück.

Räumliche Integration in der Bewegungsverarbeitung

Die Abbildung 3.35 zeigt den von Yuille und Grzywacz (1988) untersuchten Effekt der räumlichen Integration. Zusätzlich zu der in Abbildung 3.34 A dargestellten Konstellation werden ein oder mehrere Flecken zu den ursprünglichen Flecken seitlich hinzugefügt. Links hinzugefügte Flecken bewegen sich in Richtung 300° und rechts hinzugefügte in Richtung 120°. Im psychophysischen Experiment führen diese zusätzlichen Flecken zu einer eindeutigen Bewegungswahrnehmung. Die Bewegung der zusätzlichen Flecken dominiert die Bewegung der direkt benachbarten Flecken in der Wahrnehmung (*motion capture*).

Die Simulation wurde nur mit unverzögerter Linking-Kopplung und ohne Inhibition durchgeführt. Es wurden die Parameter der Tabellen 3.3 und 3.4 benutzt und die Anzahl an Flecken links und rechts von eins bis fünf variiert (Grauwert der Flecken $g_1=150$, Hintergrund $g_H=50$).

Die Simulationsergebnisse sind in Abbildung 3.36 aufgetragen. Wie bereits in Abbildung 3.34 B gezeigt, sind die Richtungsdetektoren mit Richtungspräferenzen von 0° und 120° an der Position 2 sowie die Richtungsdetektoren mit Richtungspräferenzen von 180° und 300° an der Position 3 von Interesse. Die Abbildung 3.36 A, B zeigt eine nahezu gleich starke Aktivierung der Richtungsdetektoren für 0° und 180° bei unterschiedlicher Anzahl an sich bewegenden Flecken. Beide Detektoren haben keine aktiven Detektoren mit gleicher Richtungspräferenz in ihrer Nachbarschaft im Abtastraster. Somit kommt es zu keiner Aktivitätserhöhung über Linking-Verbindungen. Sowohl der 120°- als auch der 300°-Richtungsdetektor haben hingegen aktive, benachbarte Detektoren gleicher Richtungspräferenz. Durch die laterale, unverzögerte Kopplung über Linking-Verbindungen können diese Modellneuronen ihre Aktivität um ca. 50-60% erhöhen. Mit einer Aktivitätsreduzierung durch inhibitorisch wirkende Interneuronen könnte diese höhere Aktivierung zu einer eindeutigen Richtungsdetektion an der Position 2 in Richtung 120° und an Position 3 in Richtung 300° führen, wenn an jeder Seite mindestens ein zusätzlicher Fleck hinzugefügt würde. Durch das Hinzufügen weiterer Flecken entstünden allerdings auch zusätzliche, unerwünschte Richtungsdetektorantworten. So ist beispielsweise an Position 8 (Abb. 3.35) auch der 60°-Richtungsdetektor

aktiv. Aufgrund der Richtungsdetektoraktivität ist somit nur eine Angabe der Richtung nach oben ($60^\circ/120^\circ$) oder unten ($240^\circ/300^\circ$) möglich. Weitere Mechanismen erscheinen daher notwendig, um eine eindeutige Bewegungsdetektion in realen Szenen durchführen zu können.

Zu der in Abbildung 3.34 A dargestellten Konstellation kommt in der nächsten Simulation noch ein weiteres Eingabebild mit zwei zusätzlichen Fleckenpositionen hinzu. Die Eingabe entspricht dem in Kapitel 2.1.2.2 aufgeführten psychophysischen Experiment von Anstis und Ramachandran 1987. Wie in der Abbildung 3.37 A gezeigt, bewegen sich zwei Flecken gleichen Grauwertes ($g_1=50$, $g_2=50$ und Hintergrund $g_H=100$) von den Positionen 5 und 6 auf die Positionen 1 bzw. 4. In einem weiteren Schritt findet wie zuvor ein Wechsel auf die Positionen 2 und 3 statt.

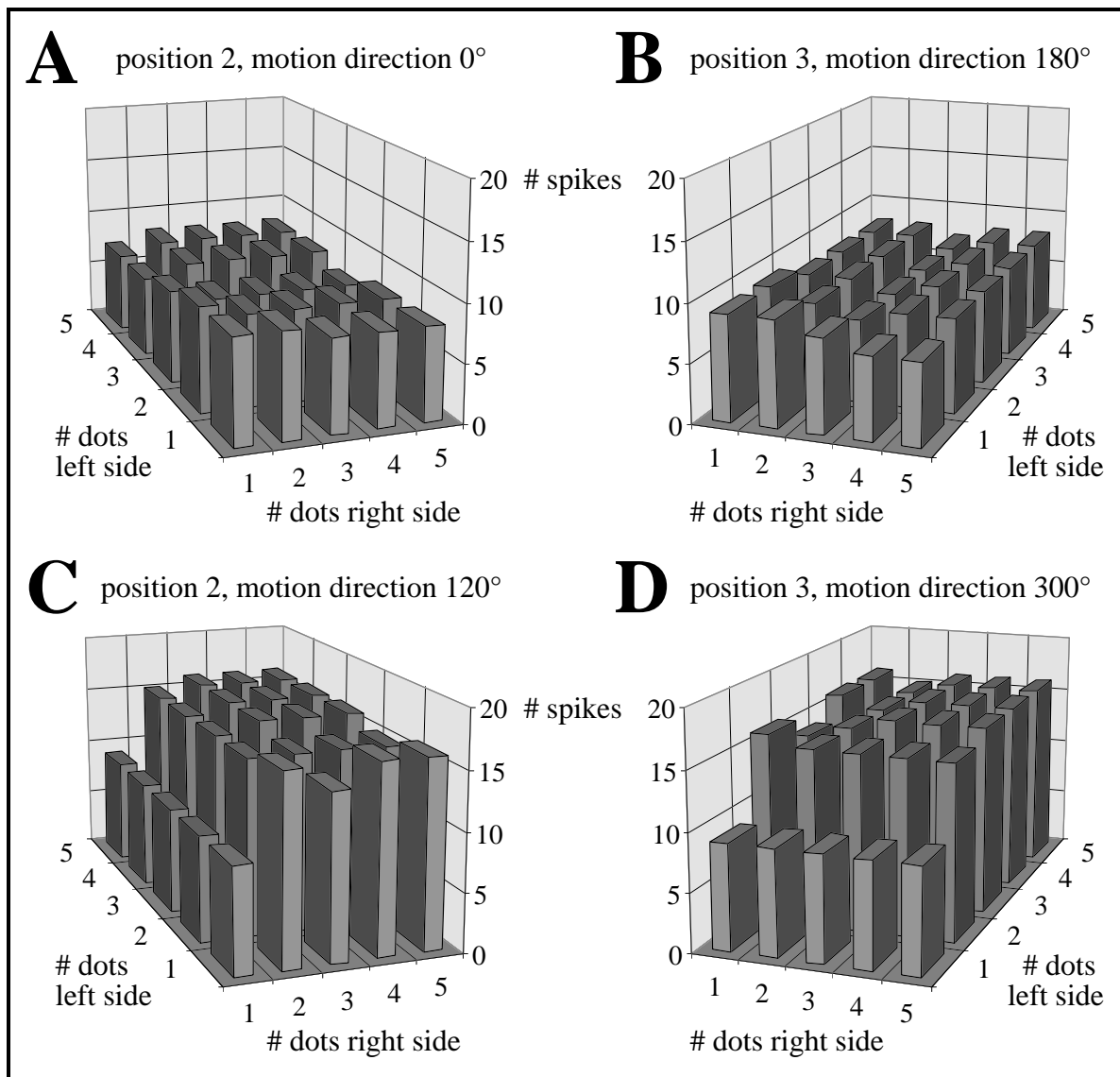


Abb. 3.36: Simulationsergebnisse zur räumlichen Integration der Bewegungsinformation. Aufgetragen ist die Anzahl an Aktionspotentialen für vier verschiedene Modellneuronen an zwei Positionen (A: Position 2, Richtungspräferenz 0° ; B: 3, 180° ; C: 2, 120° ; D: 3, 300° ; Positionen vgl. Abb. 3.30) in Abhängigkeit zur Anzahl an Flecken auf der linken (Abwärtsbewegung) und rechten (Aufwärtsbewegung) Seite. 0° - und 180° -Richtungsdetektoren erhöhen ihre Aktivität aufgrund fehlender aktiver, benachbarter Modellneuronen nicht, wohingegen 120° - und 300° -Richtungsdetektoren eine deutlich höhere Aktivität bei mindestens einem weiteren, sich bewegenden Flecken zeigen.

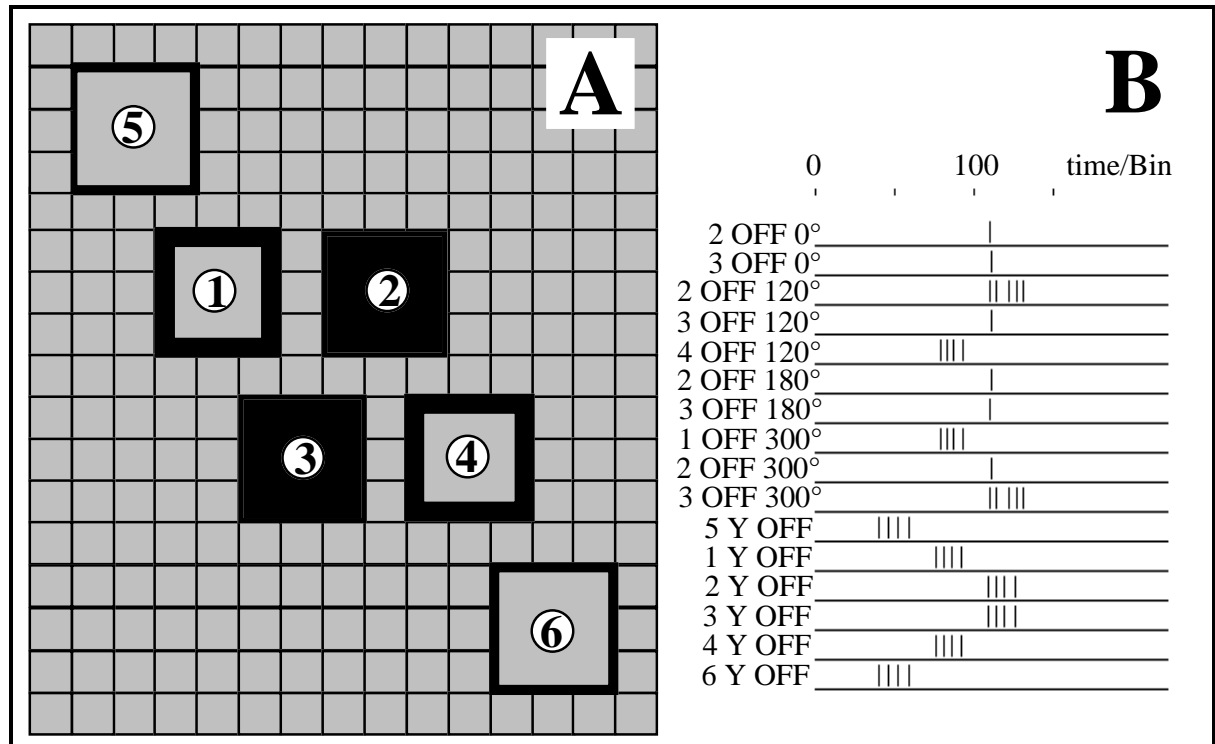


Abb. 3.37: Bewegungsdetektion mit Voraktivierung. Dargestellt ist der von Anstis und Ramachandran (1987) beschriebene „Priming-Effekt“. In **A** sind zwei Flecken aufgetragen, die sich in drei aufeinanderfolgenden Bildern von 5 nach 1 und 3 bzw. von 6 nach 2 und 4 bewegen. Zeitschritt 1: leere Vierecke mit dünnen Rahmen, Zeitschritt 2: leere Vierecke mit dicken Rahmen, Zeitschritt 3: ausgefüllte Vierecke. **B** zeigt die von den Bewegungsrichtungsdetektoren und Y-Zellen generierten Aktionspotentiale. Nähere Erläuterungen im Text.

Die Simulation wurde mit verzögertem und unverzögertem Linking sowie Inhibition durchgeführt. Es wurden die Parameter der Tabellen 3.3 und 3.4 benutzt.

Bewegungsdetektion mit Voraktivierung

Die in Abbildung 3.37 B aufgetragenen Modellneuronenaktivitäten zeigen den Einfluß der Voraktivierung. Durch die zusätzlichen Sprünge von den Positionen 6 auf 4 sowie 5 auf 1 werden Richtungsdetektoren mit RF-Mittelpunkten an den Positionen 4 bzw. 1 aktiv. Die Bewegungsdetektion ist eindeutig. Der Richtungsdetektor mit RF-Mittelpunkt an Position 1 und einer Richtungspräferenz von 300° generiert eine Aktionspotentialfolge, die an den 300°-Richtungsdetektor an Position 3 über eine Linking-Verbindung weitergegeben wird. Dadurch hat der 300°-Richtungsdetektor an dieser Position ein höheres Potential als der 180°-Richtungsdetektor an der gleichen Abtastposition. Über das inhibitorisch wirkende Interneuron kommt es zu einem lokalen Wettstreit (*winner take all*), den das stärker aktivierte Modellneuron gewinnt. Analog gilt dies für die entgegengesetzte Bewegungsrichtung an den Positionen 4 und 2.

Die Voraktivierung kann weiterhin zur Detektion eines sich geradlinig bewegenden Fleckes zwischen sich zufällig bewegenden Flecken ausgenutzt werden. Die nachfolgende Untersuchung der zeitlichen Integration der Bewegungsinformation geht auf Grzywacy et al. (1995) und Watamaniuk et al. (1995) zurück.

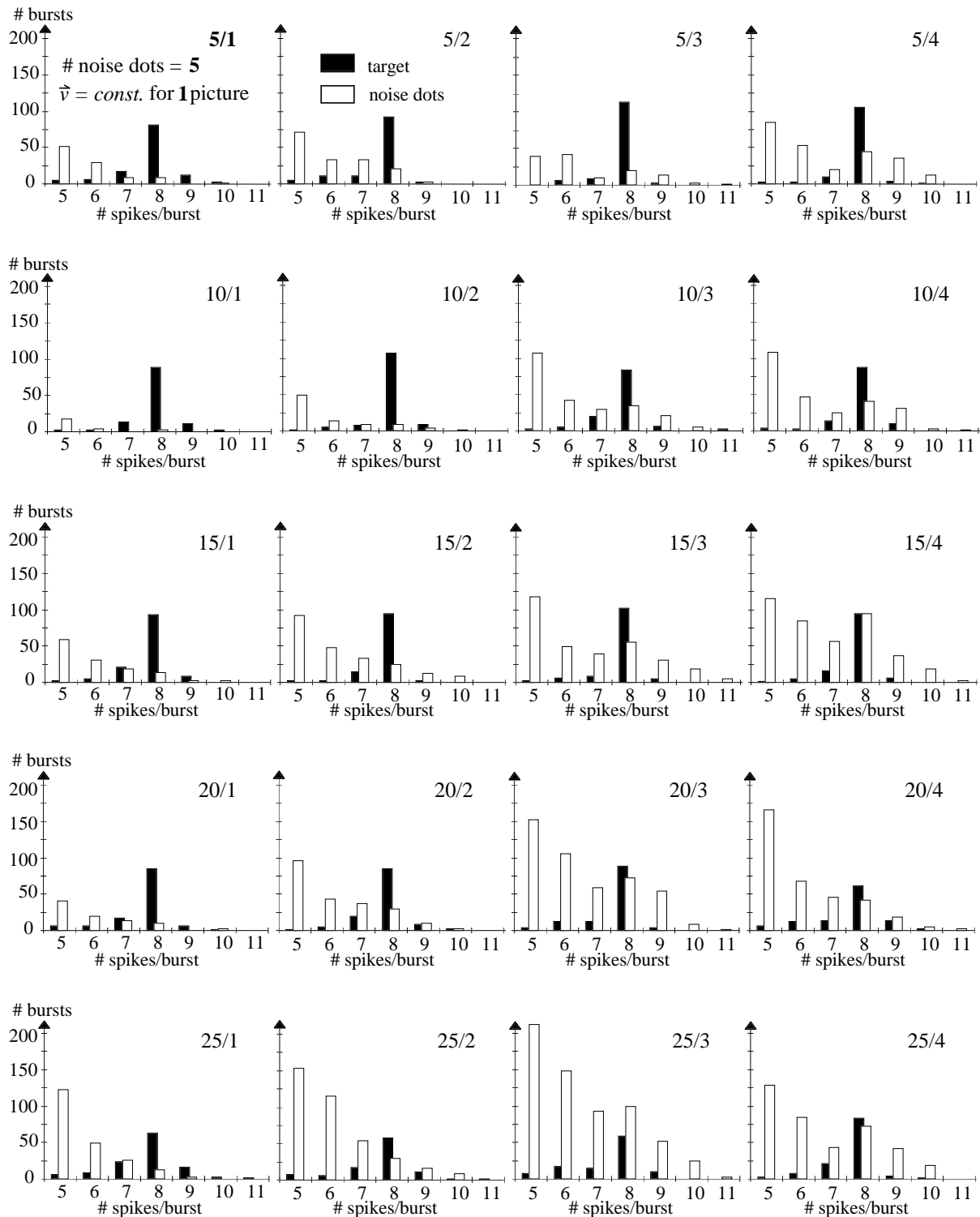


Abb. 3.38: Detektion eines geradlinig bewegten Fleckes im zufällig bewegten Fleckenmuster. Dargestellt ist die Häufigkeit (ausgefüllte Rechtecke: Zielobjekt, leere Rechtecke: Störobjekte) der Aktionspotentialanzahl pro Aktionspotentialfolge (*burst*) aller OFF-Richtungsdetektoren. Rechts oben ist jeweils die Anzahl der Störobjekte und die Anzahl an aufeinanderfolgenden Bildern mit gleicher Bewegungsrichtung eines Störobjektes aufgetragen. Nähere Erläuterungen im Text.

Als Stimulus diente ein Fleck (Zielobjekt), der sich mit einer konstanten Geschwindigkeit von $v = 4$ Pixel/Bild in Richtung 0° durch die Bildmitte bewegte und eine variierbare Anzahl (5 bis 25) an Flecken (Störobjekte), die sich ebenfalls mit $v = 4$ Pixel/Bild, aber mit stochastischer Richtungsänderung (mit gleicher Wahrscheinlichkeit in Richtung 0° , 60° , 120° , 180° , 240° und 300°) bewegten. Die Bewegungsrichtung der Störobjekte konnte über mehrere Bilder (1 bis 4) konstant gehalten werden. Alle Flecken, die wieder 3 Pixel x 3 Pixel groß waren, hatten einen Grauwert von $g_1=50$ und der Hintergrund einen Grauwert von $g_H=100$. Die Flecken konnten den Bildbereich (128 Pixel x 128 Pixel) nicht verlassen. Das Abtastgitter umfaßte $M=20 \times N=20$ Abtastpunkte. Das Zielobjekt durchlief jeweils zehnmal das Bild. Das neuronale Netz beinhaltete sowohl verzögerte und unverzögerte Linking-Verbindungen als auch eine inhibitorische Aktivitätsbeschränkung über Interneuronen.

Die Simulationsergebnisse sind in Abbildung 3.38 dargestellt. Es wurde die Anzahl an generierten Aktionspotentialen (innerhalb eines 32 ms - Zeitfensters) aller OFF-Richtungsdetektoren der mittleren Auflösung auf Stimulation mit Zielobjekt und Störobjekt getrennt bestimmt. Dabei ergab sich ein sehr hoher Anteil (bis zu 95%) an gering (≤ 4 Aktionspotentiale) aktiven Modellneuronen bei Stimulation durch Störobjekte. Die Anzahl an längeren Aktionspotentialfolgen (≥ 5 Aktionspotentiale) wurde in den Histogrammen der Abbildung 3.38 zusammengefaßt. Die Histogramme zeigen die im Vergleich zu der durch sich zufällig bewegende Flecken verursachte Aktivität starke Aktivierung durch das Zielobjekt. Die Modellneuronen, die durch das Zielobjekt stimuliert werden, generieren größtenteils acht Aktionspotentiale, die anderen deutlich weniger. Ist die Geschwindigkeit der Störobjekte nur für ein Bild konstant, so wird der sich geradlinig bewegende Fleck durch die relativ hohe Aktivität hervorgehoben. Die hohe Aktivität könnte eine Erklärung für das auffällige Hervortreten des Zielobjektes in psychophysischen Experimenten sein. Eine weitere Erhöhung der Anzahl an Störobjekten auf 25 stört die Detektion des Zielobjektes nicht. Wird jedoch zusätzlich die Bewegungsrichtung der Störobjekte über drei oder vier Bilder konstant gehalten, dann ist das Zielobjekt aufgrund seiner Aktivität nur noch bei einer geringen Anzahl an Störobjekten (< 10) von den Störobjekten separierbar. Bei Unterscheidung der Aktivität nach Richtungspräferenz der Modellneuronen wäre eine Separierung auch bei einer größeren Anzahl an Störobjekten erfolgreich, da die Aktivierung, die auf das Zielobjekt zurückgeht, sich nur auf 0° -Richtungsdetektoren beschränkt, wohingegen die Aktivierung, die durch Störobjekte verursacht wird, sich aus den Aktivitäten aller sechs Richtungsdetektorschichten zusammensetzt.

Priming-Effekt

Der „*Priming-Effekt*“ nach Pinkus und Pantle (1997) kann ebenfalls mit dem vorgestellten neuronalen Netz zur Bewegungsdetektion erklärt werden. Als Stimulus wurde ein Gitter benutzt, dessen Balken sechs Pixel breit waren und die Grauwerte $g_1=150$ und $g_2=50$ hatten (siehe Abbildung 2.5). Wird das Gitter um sechs Pixel verschoben, dann ist die Bewegungsrichtung nicht mehr eindeutig feststellbar. Erfolgt zuvor ein kleinerer Schritt, so führt eine anschließende Verschiebung um sechs Pixel bei menschlichen Beobachtern zu einer Bewegungswahrnehmung in Richtung des ersten Schrittes. In der Simulation wird das Gitter nach links oder rechts zunächst um Δ_{prime} und dann um sechs Pixel versetzt.

Die Abbildung 3.39 zeigt das Ergebnis der Simulationen. Aufgetragen wurde das Verhältnis q der zeitlich kumulierten Aktivität aller 180° ON- und OFF-Richtungsdetektoren zu der Aktivität der 0° ON- und OFF-Richtungsdetektoren gegen Δ_{prime} .

$$q = \frac{\sum_{t=50\text{Bin}}^{191\text{Bin}} (MUA_{180^\circ\text{ON}}(t) + MUA_{180^\circ\text{OFF}}(t))}{\sum_{t=50\text{Bin}}^{191\text{Bin}} (MUA_{0^\circ\text{ON}}(t) + MUA_{0^\circ\text{OFF}}(t))} \quad (\text{Gl. 3.47})$$

Die Aktivität ist hierbei ein Mittelwert aus jeweils zehn Simulationen mit zwölf verschiedenen Positionsverschiebungen des Gitters. Die Aktivität des ON- und OFF-Pfades wurde zusammengefaßt.

Ohne Verschiebung $\Delta_{\text{prime}} = 0$ Pixel oder mit einer Verschiebung von $\Delta_{\text{prime}} = 6$ Pixel sind 0°- und 180°-Richtungsdetektoren ungefähr gleich stark aktiv. Die Bewegungsrichtung kann nicht eindeutig festgestellt werden. Bei Verschiebungen $1 \text{ Pixel} \leq \Delta_{\text{prime}} \leq 5 \text{ Pixel}$ in Richtung 180° ist die 180°-Richtungsdetektoraktivität bis zu viermal größer als die der 0°-Richtungsdetektoren. Größere Verschiebungen $\Delta_{\text{prime}} \geq 7 \text{ Pixel}$ ergeben eine Scheinbewegung in Richtung 0° und führen zu einer Verstärkung der 0°-Richtungsdetektoraktivität. Für Verschiebungen in Richtung 0° gilt dies umgekehrt.

Die stärkere Aktivierung der 180° Richtungsdetektoren bei einer Verschiebung Δ_{prime} in Richtung 180° ist per se ein Resultat der Voraktivierung. Die 180°-Richtungsdetektoren werden durch die erste Verschiebung aktiv und aktivieren die in Richtung 180° benachbarten 180°-Richtungsdetektoren vor. Den 0° Richtungsdetektoren fehlt diese Voraktivierung, wodurch ihre Gesamtaktivität im Vergleich zu den 180°-Richtungsdetektoren geringer ist.

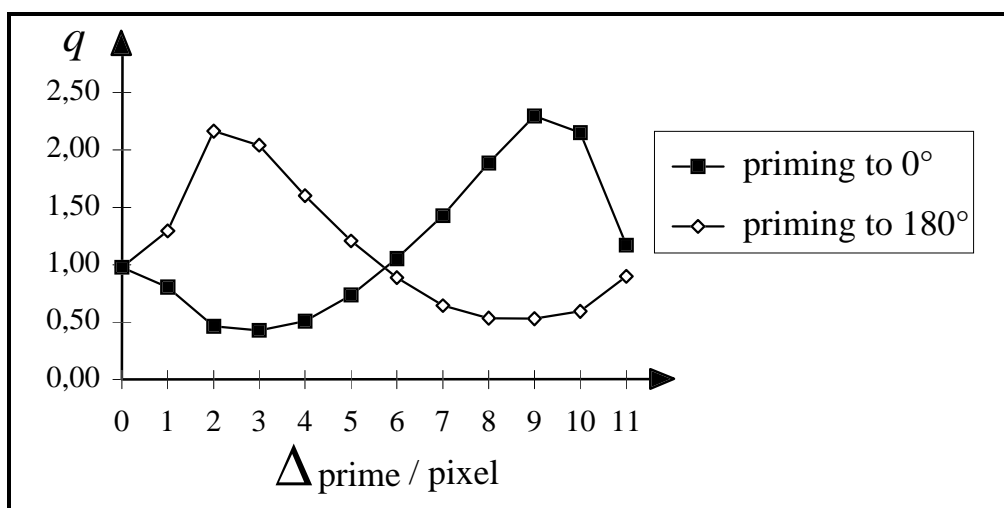


Abb. 3.39: Simulationsergebnis „Priming“. Aufgetragen ist das Verhältnis der zeitlich und räumlich kumulierten Aktivitäten der 180°- zu den 0°-Richtungsdetektoren. Das Simulationsergebnis zeigt den von Pinkus und Pantle (1997) beschriebenen Effekt. Ein Gitterreiz mit einer Balkenbreite von 6 Pixel wird zunächst um Δ_{prime} in Richtung 0° (ausgefüllte Rechtecke) bzw. 180° (leere Rauten) versetzt und anschließend in die gleiche Richtung um 6 Pixel verschoben. Die Bewegungsrichtung der zweiten Verschiebung ist somit nicht feststellbar. Durch die eindeutige Bewegungsbestimmung des ersten Sprunges kommt es zu einer eindeutigen Bewegungsrichtungsdetektion der zweiten Verschiebung.

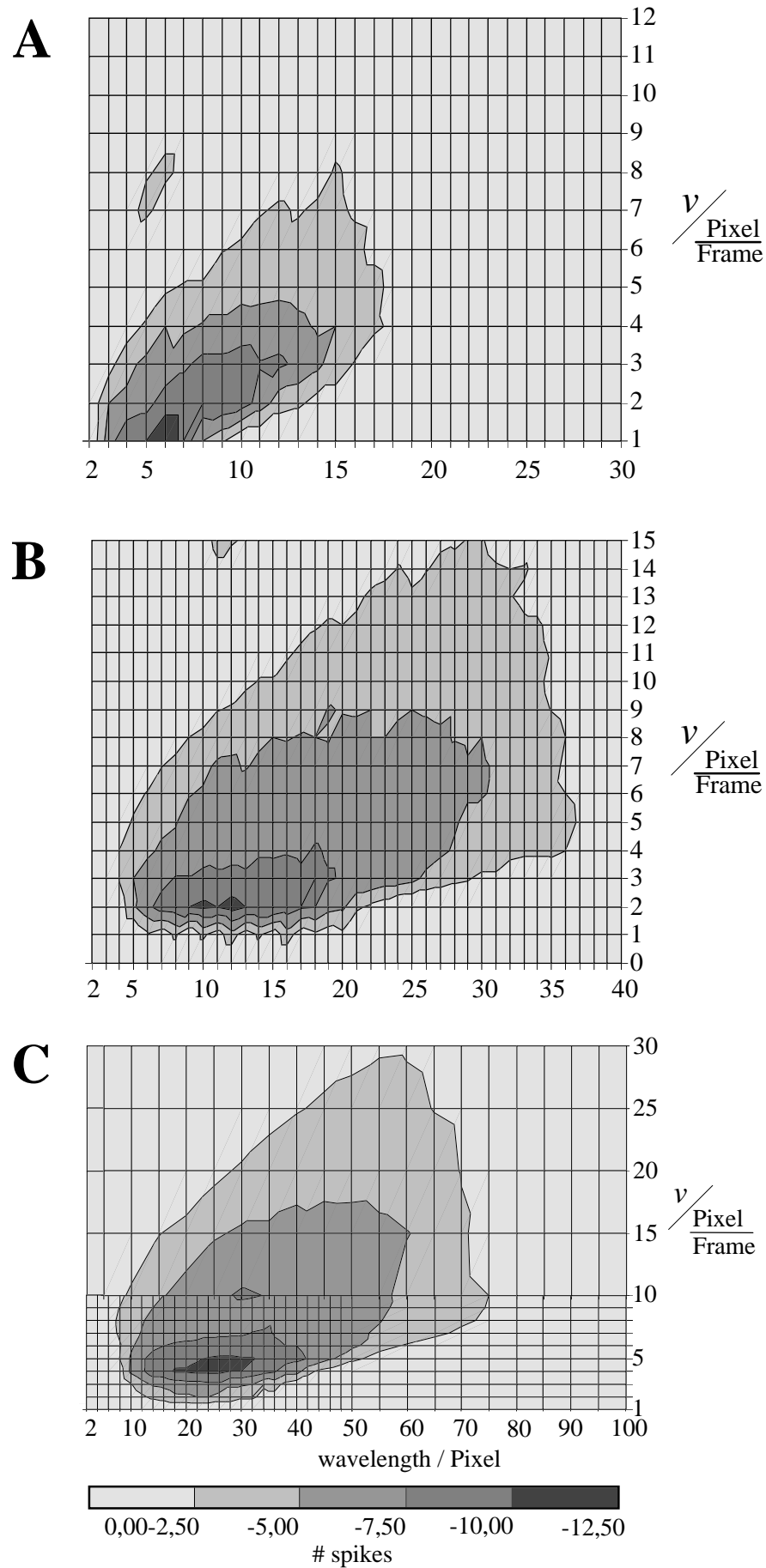


Abb. 3.40: Geschwindigkeitsabhängigkeit der 180°-Richtungsdetektoren. A: feine, **B:** mittlere und **C:** grobe Auflösung. Aus Gründen der Darstellung sind die Achsen unterschiedlich skaliert. Erläuterungen im Text.

Kantenschärfe

Wichtige Hinweise auf Korrespondenz geben räumliche Nähe und räumliche Auflösung. In dem vorgestellten Modell zur Bewegungsdetektion wird eine Ausnutzung dieser Hinweise durch Richtungsdetektoren mit drei unterschiedlichen räumlichen Auflösungen ermöglicht. Zur Untersuchung der Richtungsdetektoraktivität in Abhängigkeit zum Intensitätsgradienten wurde die bereits in Kapitel 3.1.2.1 benutzte Einzelwelle mit einer eindimensionalen sinusförmigen Intensitätsverteilung verwendet. Die Welle wurde in Richtung 180° mit unterschiedlichen Geschwindigkeiten v bewegt. Zusätzlich wurde die Wellenlänge variiert.

Die Abbildung 3.40 zeigt die Mittelwerte der Anzahl an Aktionspotentialen in einer Grauwertdarstellung für jeweils einen 180° -OFF-Richtungsdetektor aus der Mitte des Abtastrasters der feinen (A), mittleren (B) und groben (C) Auflösung. Die Abbildung stellt zum einen die Abhängigkeit der Aktivierung von der Geschwindigkeit und zum anderen die Abhängigkeit von der Wellenlänge der Sinusschwingung dar. Die Richtungsdetektoren der feinen, mittleren und groben Auflösung sind bei Stimulusgeschwindigkeiten von $v = 1$ Pixel/Bild, $v = 2$ Pixel/Bild bzw. $v = 4$ Pixel/Bild bei Wellenlängenbereichen von 5 bis 7 Pixel, 9 bis 13 Pixel bzw. 18 bis 32 Pixel am stärksten aktiv. Die Geschwindigkeitsbereiche sind abhängig von der Art des Stimulus.

Das Simulationsergebnis zeigt, daß die Richtungsdetektoren unterschiedlicher Auflösung auf unterschiedlich scharfe Kanten mit unterschiedlicher Geschwindigkeit verschieden stark antworten. Präzise lokalisierte Kanten werden in der feinen Auflösung, Kanten mit einem flachen Intensitätsgradienten in der mittleren oder groben Auflösung verarbeitet. Durch diese parallele Verarbeitung wird das Korrespondenzproblem reduziert. Das gleiche gilt für die Geschwindigkeit. Kleine Sprünge werden von den Richtungsdetektoren der feinen Auflösung und größere Sprünge von den Richtungsdetektoren der mittleren oder groben Auflösung detektiert. Somit können die Hinweise „Nähe“ und „Schärfe“ (die räumlichen Frequenzanteile) für die Bewegungsrichtungsdetektion ausgenutzt werden.

Richtungscharakteristiken

In der letzten Simulation wurde nur die Aktivität eines 180° -Richtungsdetektors bei einer Kantenbewegungsrichtung betrachtet. In den nachfolgenden Simulationen werden die Richtungscharakteristiken der Richtungsdetektoren untersucht. Als Stimulus dienten zunächst Kanten unterschiedlicher Orientierung, die mit Geschwindigkeiten von $v = 4$ Pixel/Bild orthogonal zur Kantenorientierung über das Abtastraster bewegt wurden. Die Grauwertdifferenzen der Kanten $\Delta g = g_1 - g_2$ (mit $g_2=50$) wurden in den Simulationen variiert, um die Abhängigkeit der Richtungscharakteristiken von der Grauwertdifferenz zu bestimmen. Die Eingabebilder waren 128 Pixel x 128 Pixel groß und rauschfrei. Das Abtastraster befand sich in der Mitte der Eingabebilder und war kreisförmig (Radius 20 Pixel), damit die Kanten aller Orientierungen gleich lang waren. Die Simulationen wurden zum einen ohne und zum anderen mit lateralen Kopplungen - dazu gehören verzögerte und unverzögerte Linking-Verbindungen sowie die Inhibition - der Richtungsdetektoren durchgeführt. Es wurden die Parameter der Tabellen 3.3 und 3.4 verwendet.

Die mittleren Aktivierungen der ON-Richtungsdetektoren (Abtastweite $d = 4$ Pixel) an zwölf Abtastorten in der Mitte des Abtastrasters sind in den Abbildungen 3.41 und 3.42

aufgetragen. Die Abbildung 3.41 zeigt, daß die Richtungsdetektoren bereits ohne laterale Kopplungen richtungsspezifisch antworten. Die Anzahl der Aktionspotentiale pro Reizdurchgang wird hierbei mit zunehmender Grauwertdifferenz größer. Allerdings ändern sich auch die Richtungscharakteristiken der Richtungsdetektoren. Mit zunehmender Grauwertdifferenz überlappen sich die Aktivitätsverteilungen etwas stärker. Dies wird besonders in der Abbildung 3.43 A deutlich. Diese Abbildung zeigt die Halbwertsbreite der Richtungscharakteristiken der 0° -Richtungsdetektoren in Abhängigkeit zur Grauwertdifferenz. Die Halbwertsbreite wurde dabei aus der Differenz der Kantenorientierungen, die zu einer halb so großen Aktivität der 0° -Richtungsdetektoren wie eine 90° orientierte Kante führen, bestimmt. Die große Überlappung der Aktivitätsverteilungen, d.h. die Mehrdeutigkeit des Bewegungsanalyseergebnisses, resultiert aus dem Korrespondenzproblem. So ist ein 0° -Richtungsdetektor auch bei Stimulation mit einer 300° oder 60° orientierten Kante aktiv.

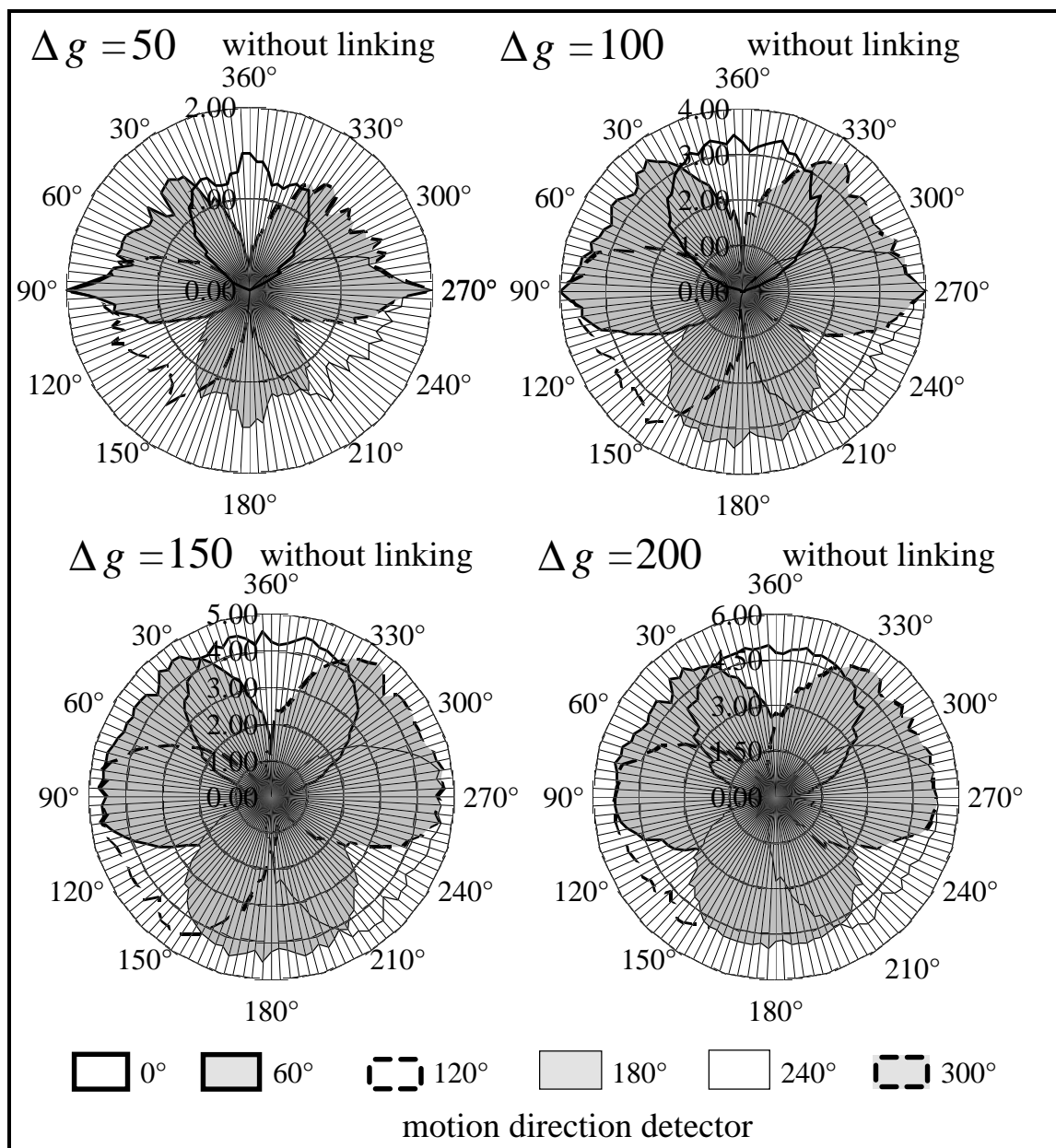


Abb. 3.41: Richtungscharakteristiken ohne Linking-Kopplung. Für vier verschiedene Grauwertdifferenzen Δg sind die mittleren Impulsaktivitäten der ON-Richtungsdetektoren der mittleren Auflösung ohne Linking-Verbindungen bei Stimulation mit einer sich bewegenden Kante ($v = 4$ Pixel/Bild) in Abhängigkeit zur Bewegungsrichtung radial aufgetragen. Die Richtungscharakteristiken überlappen sich mit zunehmender Grauwertdifferenz immer mehr.

Die Abbildung 3.42 zeigt die Richtungscharakteristiken der Richtungsdetektoren mit lateraler Kopplung über Linking-Verbindungen. Das Ergebnis der Bewegungsanalyse ist aufgrund des lokalen Wettstreits eindeutig. Die Richtungscharakteristiken sind scharf und überlappen sich nur bei Kantenbewegungsrichtungen, die genau zwischen zwei Richtungspräferenzen liegen. Ferner sind die Breiten der Aktivitätsverteilungen für alle Grauwertdifferenzen gleich (siehe auch Abbildung 3.43 A). Die Aktivität der Bewegungsrichtungsdetektoren ist auch hier von der Grauwertdifferenz abhängig. Die Anisotropie in der Gesamtaktivitätsverteilung, insbesondere in Richtung 90° und 270°, ist auf eine verschiedene Anzahl an aktiven Modellneuronen zurückzuführen, die eine unterschiedliche Erhöhung der Aktivität über die lateralen Kopplungen bewirken.

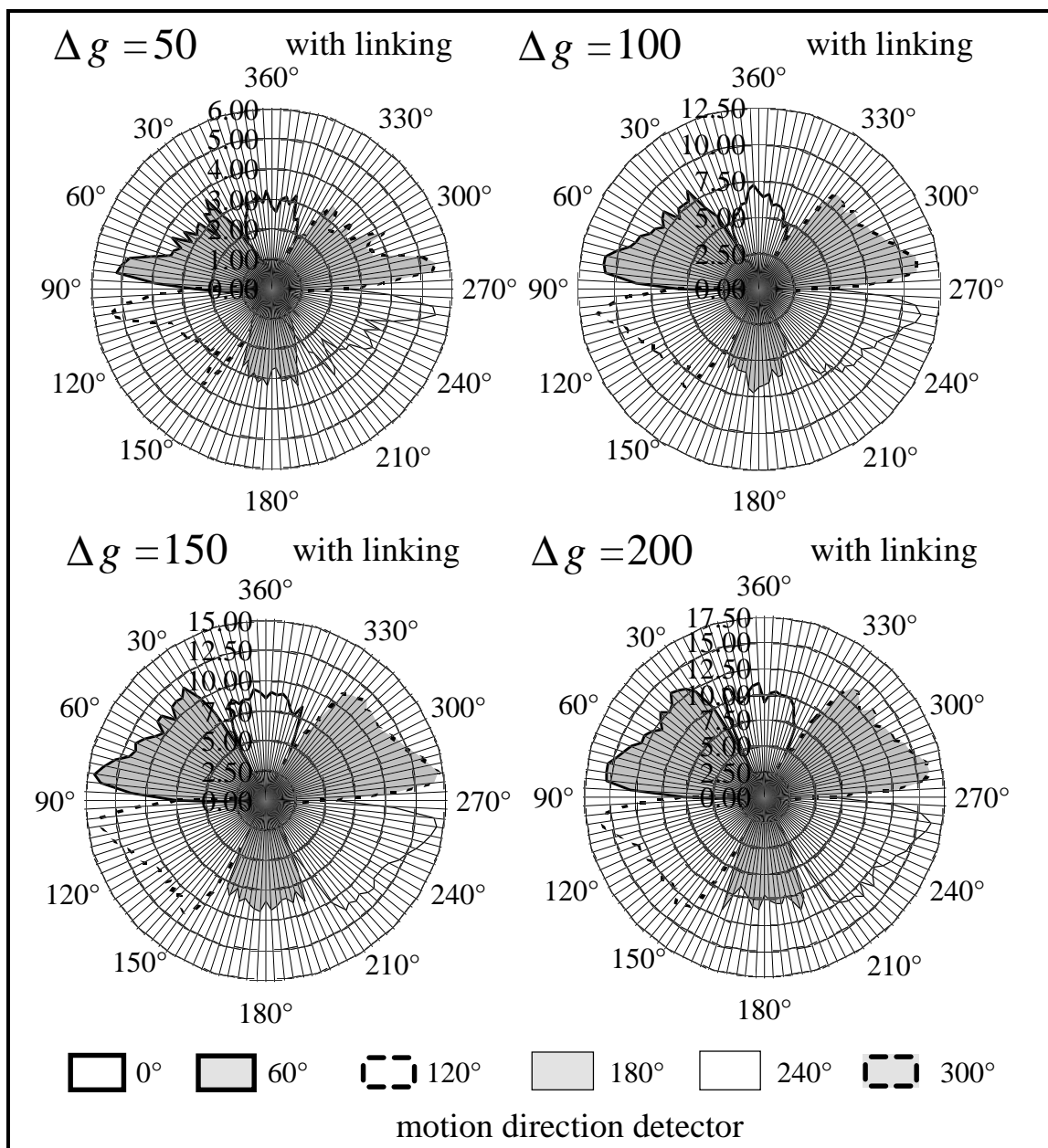


Abb. 3.42: Richtungscharakteristiken der Richtungsdetektoren mit Linking-Kopplung. Für vier verschiedene Grauwertdifferenzen Δg sind radial die mittleren Aktivitäten der ON-Richtungsdetektoren der mittleren Auflösung mit Linking-Verbindungen auf Stimulation mit einer sich bewegenden Kante ($v = 4$ Pixel/Bild) in Abhängigkeit zur Bewegungsrichtung aufgetragen. Die Breiten der Aktivitätsverteilungen sind unabhängig von der Grauwertdifferenz.

Der Vergleich der Abbildungen 3.41 und 3.42 zeigt auch, daß das Ergebnis der Richtungsdetektion mit lateralen Kopplungen eindeutig wird und die Halbwertsbreiten der Aktivitätsverteilungen unabhängig von der Grauwertdifferenz der Kante sind. Bei Beachtung der Gesamtaktivität, das ist in diesem Fall die Summe der mittleren Aktivitäten über alle Kantenbewegungsrichtungen, läßt sich eine Umverteilung der Aktivität feststellen. Bei einer hohen Grauwertdifferenz von $\Delta g = 200$ ist die Gesamtaktivität der Richtungsdetektoren mit und ohne Linking-Verbindungen etwa gleich groß. Die Richtungscharakteristiken der Richtungsdetektoren mit lateraler Kopplung sind schärfer (Abbildung 3.43 A) und ihre Aktivitäten höher (Abbildung 3.43 B) als die der Richtungsdetektoren ohne laterale Kopplung. Bei einer geringen Grauwertdifferenz von $\Delta g = 50$ ist die Gesamtaktivität mit lateralen Kopplungen ungefähr doppelt so groß wie die Gesamtaktivität ohne laterale Kopplungen. Dadurch hebt sich die von der Kante verursachte neuronale Aktivität besser von der durch das Rauschen verursachten ab (siehe unten). Die Linking-Verbindungen führen in Hinblick auf Aktivierung und Eindeutigkeit zu einer besseren Detektion der Kantenbewegungsrichtung.

In der Abbildung 3.40 wurde bereits gezeigt, daß die 180° -Richtungsdetektoren entsprechend der Größe ihrer rezeptiven Felder Kanten mit einem bestimmten Geschwindigkeitsbereich bevorzugt detektieren können. In den drei nachfolgenden Abbildungen 3.44, 3.45 und 3.46 sind die Richtungscharakteristiken der lateral über Linking-Verbindungen gekoppelten Modellneuronen für verschiedene Kantengeschwindigkeiten aufgetragen. Die Grauwertdifferenz der Kante betrug $\Delta g = 100$. Die Auswertung erfolgte wie bereits oben beschrieben.

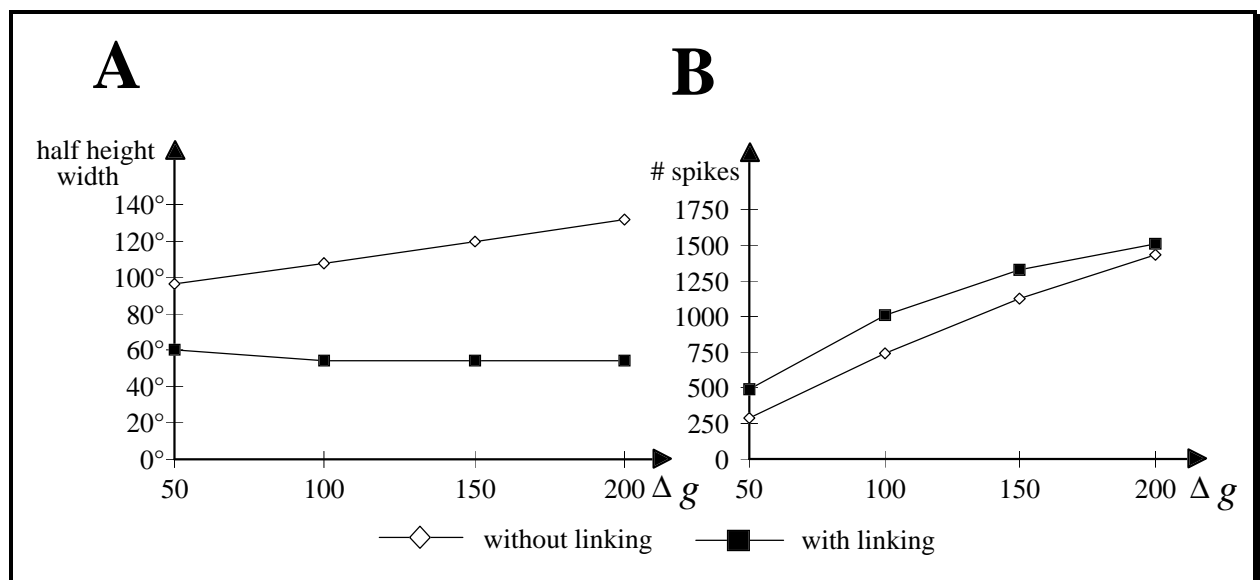


Abb. 3.43: Breite der Richtungscharakteristiken und Gesamtaktivität der Richtungsdetektoren.

A: Ausgehend von den Abbildungen 3.35 und 3.36 wurde die Halbwertsbreite der Richtungscharakteristiken der 0° -Richtungsdetektoren in Abhängigkeit zur Grauwertdifferenz Δg aufgetragen. Ohne laterale Linking-Kopplungen werden die Richtungscharakteristiken mit zunehmender Grauwertdifferenz breiter. Die Breite der Richtungscharakteristik ist bei Richtungsdetektoren mit lateralen Linking-Kopplungen nahezu unabhängig von der Grauwertdifferenz. **B:** Die aufsummierte Gesamtaktivität der 0° -Richtungsdetektoren zeigt, daß bei einer kleinen Grauwertdifferenz $\Delta g = 50$ die lateral gekoppelten Modellneuronen doppelt so stark aktiv sind wie die nicht gekoppelten. Die Gesamtaktivität bei großen Grauwertdifferenzen ist hingegen nahezu gleich.

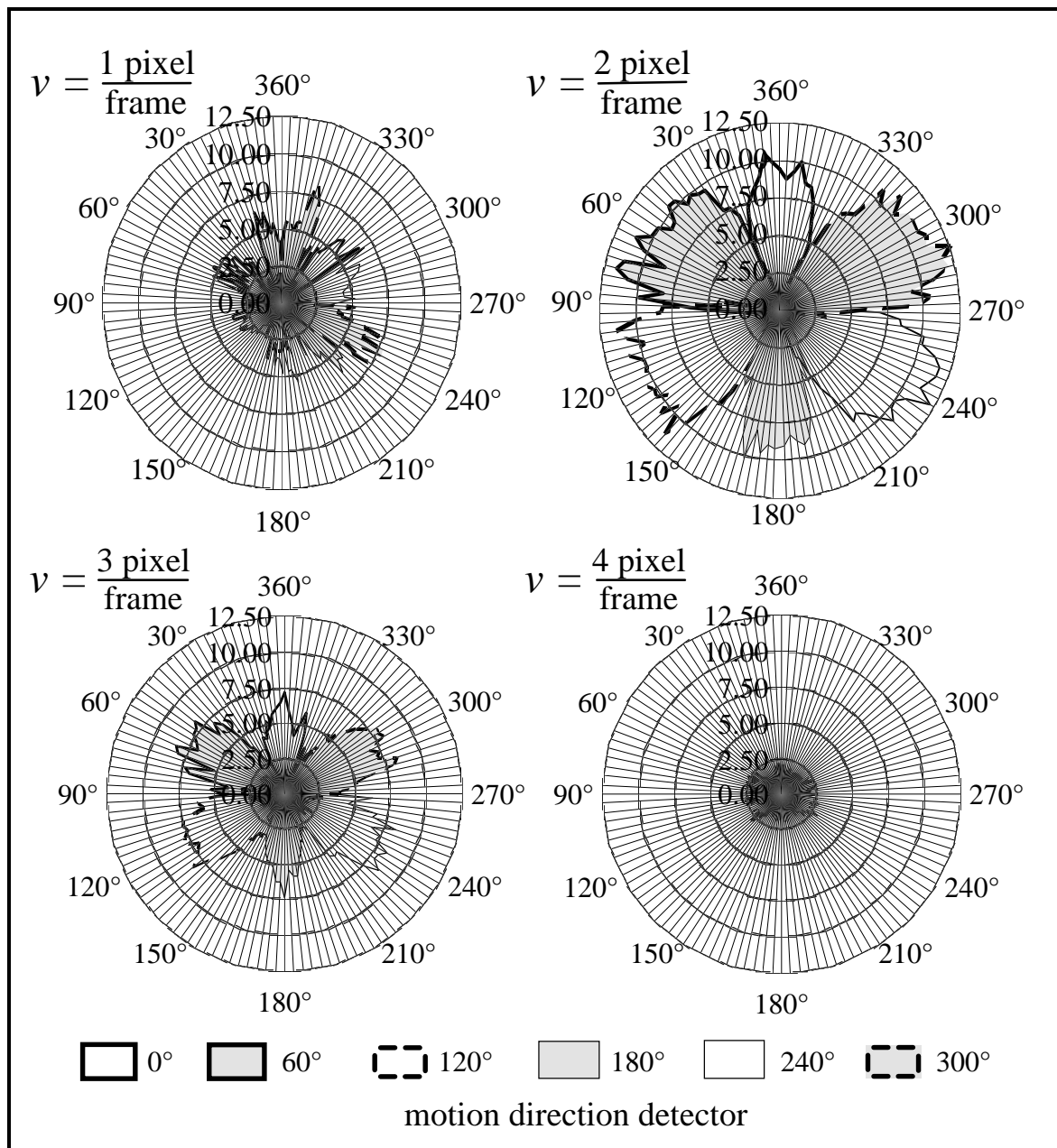


Abb. 3.44: Richtungscharakteristiken der ON-Richtungsdetektoren der feinen Auflösung.

Radial ist die mittlere Aktivität der sechs ON-Richtungsdetektoren der feinen Auflösung in Abhängigkeit von der Kantenbewegungsrichtung für vier verschiedene Geschwindigkeiten aufgetragen. Die Richtungsdetektoren sind lateral über Linking-Verbindungen gekoppelt. Die Grauwertdifferenz der Kante betrug 100. Die Richtungsdetektoren der feinen Auflösung sind besonders sensitiv für Kantengeschwindigkeiten von $v = 2$ Pixel/Bild und $v = 3$ Pixel/Bild.

In den Abbildungen 3.44 und 3.45 sind die Richtungscharakteristiken der ON- bzw. OFF-Richtungsdetektoren der feinen Auflösung dargestellt. Auf Stimulation mit einer hell/dunkel-Kante antworten per se die Richtungsdetektoren des ON-Pfades stärker als die des OFF-Pfades. Die Modellneuronen beider Pfade detektieren Kanten mit einer Geschwindigkeit von $v = 2$ Pixel/Bild am besten. Sowohl bei den ON- als auch bei den OFF-Richtungsdetektoren ist eine scharfe Richtungscharakteristik zu erkennen. Das Ergebnis der Richtungsdetektion ist eindeutig.

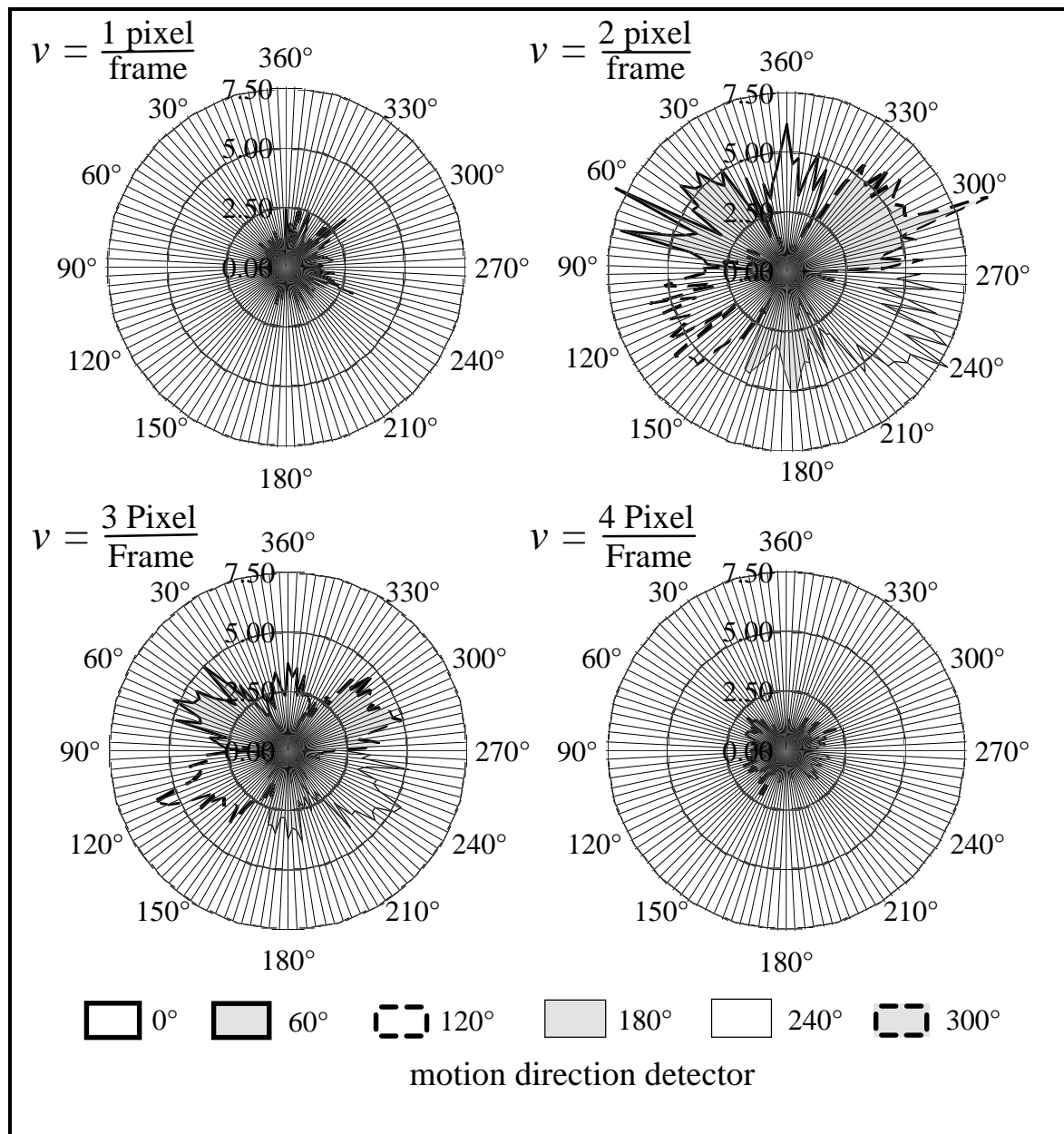


Abb. 3.45: Richtungscharakteristiken der OFF-Richtungsdetektoren der feinen Auflösung.

Radial ist die mittlere Aktivität der sechs OFF-Richtungsdetektoren der feinen Auflösung in Abhängigkeit von der Kantenbewegungsrichtung für vier verschiedene Geschwindigkeiten aufgetragen. Auch die OFF-Richtungsdetektoren der feinen Auflösung sind sensitiv für Kantengeschwindigkeiten von $v = 2$ Pixel/Bild und $v = 3$ Pixel/Bild, allerdings zeigen sie im Vergleich zu den ON-Richtungsdetektoren eine geringere Aktivität.

Die Richtungscharakteristiken der ON-Richtungsdetektoren der mittleren Auflösung sind in der Abbildung 3.46 aufgetragen. Das Ergebnis der Bewegungsdetektion ist für kleine Geschwindigkeiten nicht eindeutig. Erst ab einer Kantengeschwindigkeit von $v = 4$ Pixel/Bild bilden sich zusammenhängende und nicht überlappende Richtungscharakteristiken heraus. Die Mehrdeutigkeit der Bewegungsdetektion bei kleinen Geschwindigkeiten kann, wie in Kapitel 3.3 gezeigt wird, durch die Kopplung der Richtungsdetektoren unterschiedlicher Auflösungen beseitigt werden. Die Richtungsdetektoren des OFF-Pfades weisen wiederum ähnliche Richtungscharakteristiken, aber mit kleineren Aktivitätswerten, auf.

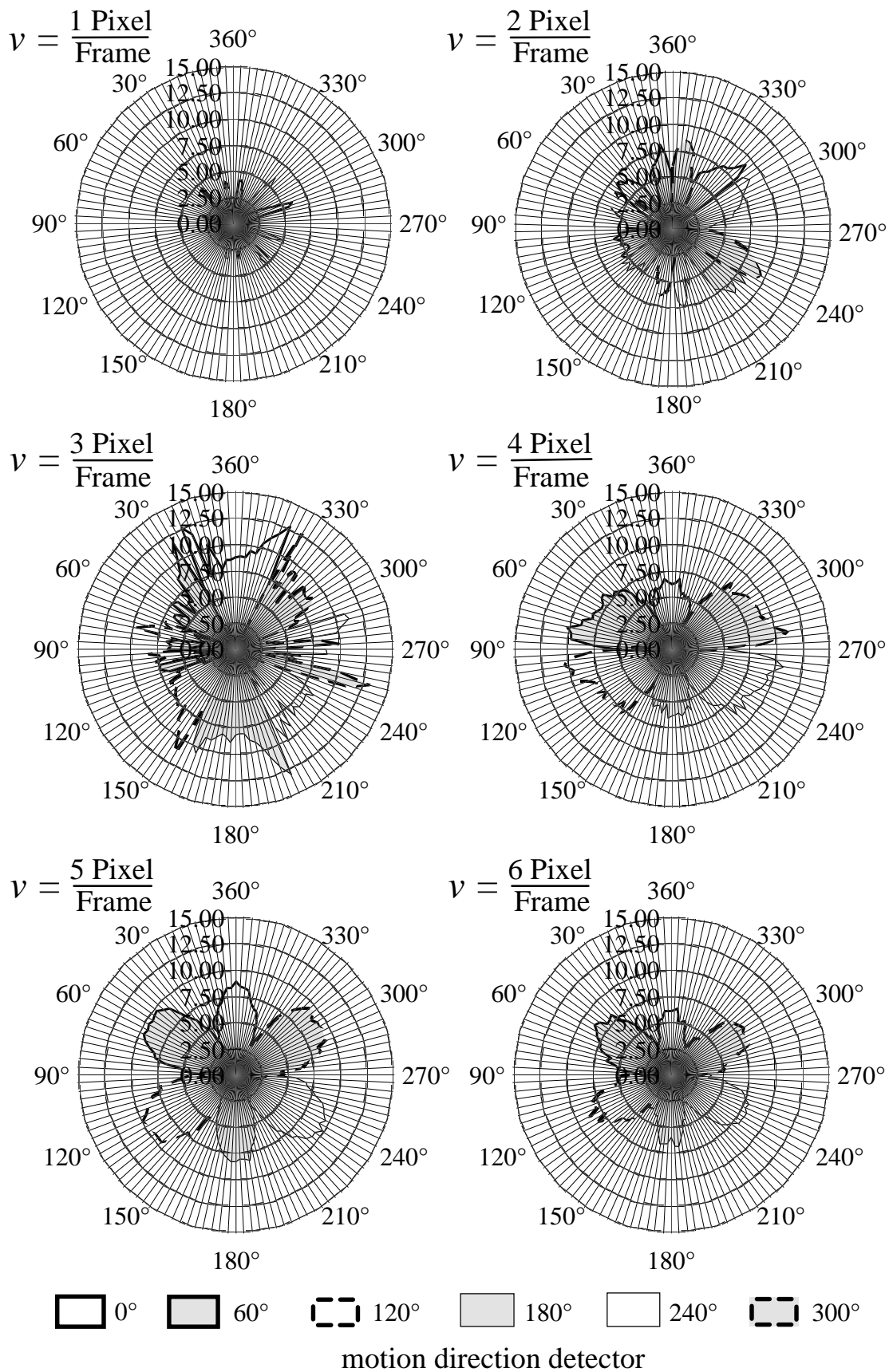


Abb. 3.46: Richtungscharakteristiken der ON-Richtungsdetektoren der mittleren Auflösung. Radial ist die mittlere Aktivität der sechs ON-Richtungsdetektoren in Abhängigkeit von der Kantenbewegungsrichtung für sechs verschiedene Geschwindigkeiten aufgetragen. Ab einer Geschwindigkeit von $v = 4 \text{ Pixel/Bild}$ ist die Bewegungsdetektion eindeutig.

Eingangsbilder mit Rauschüberlagerung

In den nächsten Simulationen wird das Antwortverhalten der Richtungsdetektoren auf verrauschte Eingabebilder (z.B. Kamerarauschen bei schlechten Bildverhältnissen) genauer untersucht. Hierbei wird sich zeigen, daß die Verwendung von Richtungsdetektoren, insbesondere mit lateraler Verschaltung über Linking-Verbindungen, zur Erhöhung des Signal/Rausch-Verhältnisses führen.

Als Eingabe diente in diesen Untersuchungen eine 90° orientierte Intensitätskontrastkante mit den Grauwerten $g_1=160$ und $g_2=90$, die sich in Richtung 0° mit einer Geschwindigkeit von $v = 2$ Pixeln/Bild ($\Delta=0$ Pixel) bewegte. Eine Sequenz umfaßte 60 Bilder. Zunächst wurden alle Pixel der 128 Pixel x 128 Pixel großen Eingabebilder zeitlich verrauscht. Dazu wurde zu dem Grauwert jedes Pixels ein Rauschwert addiert. Die Rauschwerte waren gleichverteilt aus einem Intervall $[-V^{\text{noise}}, +V^{\text{noise}}]$ (V^{noise} : Rauschamplitude), was näherungsweise zu räumlichem weißen Rauschen in den Eingabebildern führte. In weiteren Simulationen wurden die Grauwerte von jeweils vier direkt benachbarten Pixeln mit dem gleichen Rauschwert verrauscht. Des weiteren wurden dem neuronalen Netzwerk Eingabebilder, die nur Rauschen (Grauwert $g_1=g_2=125$) und keine Kante beinhalteten, angeboten.

Die Abbildung 3.47 zeigt die kumulierten Aktivitäten der ON-Richtungsdetektoren der feinen Auflösung ($d = 2$ Pixel) mit und ohne laterale Linking-Verbindungen in Abhängigkeit von der Rauschamplitude V^{noise} . Die von den Richtungsdetektoren generierten Aktionspotentiale bis zum Zeitschritt $t = 63$ Bin bleiben wegen des Anschalteffekts unberücksichtigt. Das Abtastraster bestand aus 60 Abtastpunkten pro Zeile und 20 Abtastpunkten pro Spalte.

Ohne laterale Linking-Verbindungen (Abbildung 3.47 A, B) ist das Ergebnis der Bewegungsrichtungsdetektion nicht eindeutig. Die 0° -Richtungsdetektoren sind zwar am aktivsten, doch auch die 60° - und 300° -Richtungsdetektoren zeigen eine nicht vernachlässigbare Aktivität. Mit Linking-Verbindungen werden die 60° - und 300° -Richtungsdetektoren unterdrückt und die Bewegungsdetektion wird somit eindeutig (Abbildung 3.47 C, D). Die 120° -, 180° - und 240° -Richtungsdetektoren sind bei kleinen Rauschamplituden noch nicht aktiv. Steigt die Rauschamplitude, so steigen auch deren Aktivitäten. Ihre Aktivität resultiert, wie ein Vergleich mit der Simulation zeigt, in der dem Netz nur Rauschen angeboten wurde, nicht aus der Intensitätskontrastkante, sondern ausschließlich aus dem Hintergrundrauschen. Der Aktivitätsverlauf der 0° -Richtungsdetektoren fällt zunächst mit steigender Rauschamplitude ($V^{\text{noise}} \leq 50$) ab und steigt dann erst an. Der Aktivitätsabfall ist auf die durch das Rauschen reduzierte Aktivität der Richtungsdetektoren, die unmittelbar von der Kante gereizt werden, zurückzuführen. Der Anstieg des Aktivitätsverlaufs resultiert aus dem Rauschen des Hintergrundes. Das Rauschen im Hintergrund ist bei kleinen Rauschamplituden ($V^{\text{noise}} \leq 30$) nicht ausreichend, um Membranpotentiale zu generieren, die den Schwellenoffset der Modellneuronen überschreiten, wohingegen sich bereits schwaches Rauschen ($V^{\text{noise}} = 10$) an der Kante unmittelbar in der Aktionspotentialgenerierung bemerkbar macht.

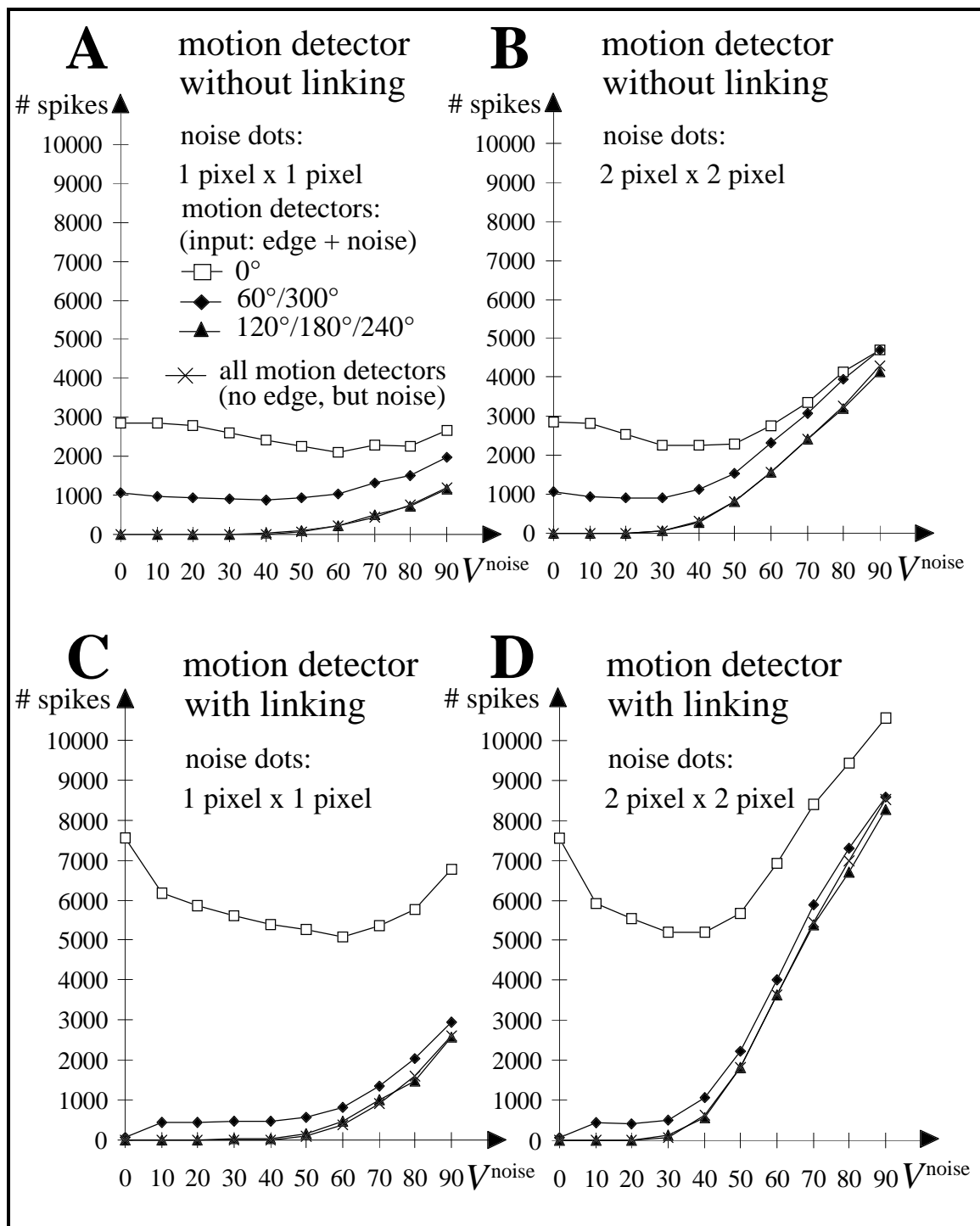


Abb. 3.47: Rauschabhängigkeit der Richtungsdetektoraktivität. In der Abbildung sind die kumulierten Aktivitäten der ON-Richtungsdetektoren in Abhängigkeit von der Rauschamplitude V^{noise} dargestellt. Als Stimulus dienten eine sich in Richtung 0° bewegende, 90° orientierte verrauschte Kante und Rauschen. Die Aktivitätswerte, die ausschließlich auf Rauschen beruhen, sind Mittelwerte aller Richtungsdetektoren. Die Simulationen wurden mit (C, D) und ohne (A, B) laterale Linking-Verbindungen durchgeführt. In A, C wurde jeder einzelne Pixel zeitlich verrauscht, in B, D wurden jeweils vier benachbarte Pixel mit demselben Rauschwert verrauscht. Ohne Linking-Interaktionen kann bei starkem Rauschen anhand der Richtungsdetektoraktivität keine Kantenbewegungsrichtung mehr bestimmt werden.

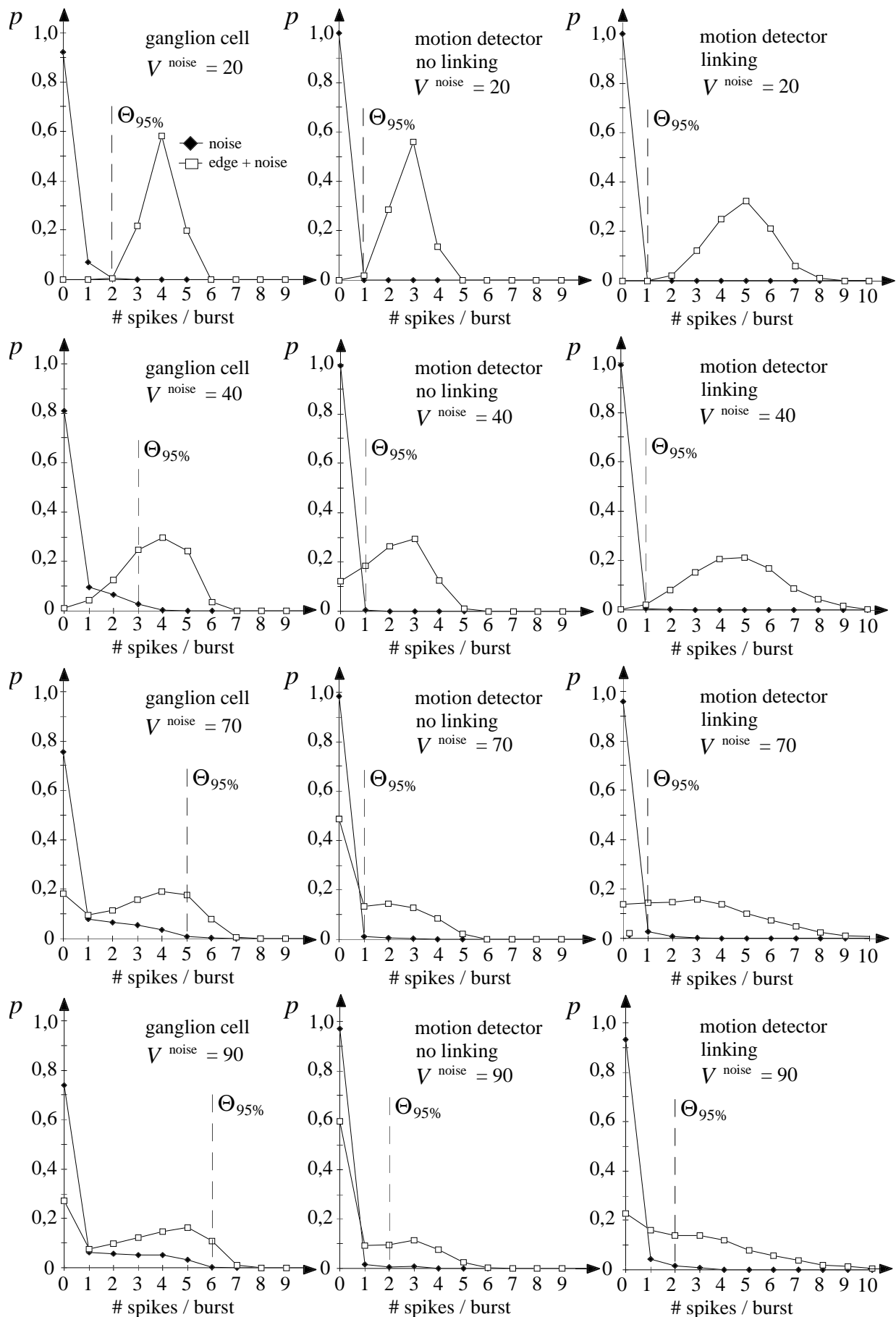


Abb. 3.48: Relative Häufigkeiten der Aktionspotentialanzahl von retinalen Ganglienzellen und Richtungsdetektoren bei Stimulation mit einer verrauschten Kante. Ausführliche Erläuterungen im Text.

Aus der kumulierten Aktivität der Richtungsdetektoren läßt sich noch nicht vorhersagen, inwieweit nachgeschaltete Verarbeitungseinheiten die Kantenbewegungsrichtung aus dem Rauschen extrahieren können. Von Bedeutung sind hierbei wegen der begrenzten zeitlichen Integrationsfähigkeit der nachfolgenden Modellneuronen die kurzzeitigen Aktivitäten der Richtungsdetektoren. Die Entscheidung, ob ein Signal oder nur Rauschen bei Betrachtung eines einzelnen Richtungsdetektors vorliegt, läßt sich nur anhand der Anzahl an generierten Aktionspotentialen in dessen Aktionspotentialfolge treffen. Je größer diese Anzahl ist, desto wahrscheinlicher ist es, daß der Richtungsdetektor von einer Intensitätskontrastkante stimuliert wurde. In der Abbildung 3.48 sind die relativen Häufigkeiten p für eine bestimmte Anzahl an Aktionspotentialen pro Aktionspotentialfolge bei Reizung mit einer verrauschten Kante und reinem Rauschen unter Variation der Rauschamplitude V^{noise} aufgetragen. Hierbei wurden sowohl 0° -Richtungsdetektoren mit und ohne laterale Linking-Verbindungen als auch retinale Ganglienzellen in die Untersuchung miteinbezogen. Zur Berechnung der relativen Häufigkeit wurden jeweils die 40 mittleren Abtastorte jeder zweiten Zeile herangezogen, so daß sich ohne Berücksichtigung der letzten Zeile und bei zehn Simulationen 3600 ausgewertete Aktionspotentialfolgen ergeben. Die Länge des Zeitfensters zur Bestimmung der Aktionspotentialanzahl pro Aktionspotentialfolge wurde auf 32 Bin festgelegt.

Die Abbildung 3.48 zeigt eine Verbreiterung der Verteilung der relativen Häufigkeit bei zunehmender Rauschamplitude. Bei den retinalen Ganglienzellen ist dies besonders stark ausgeprägt. Die Wahrscheinlichkeit, daß eine retinale Ganglienzelle auf Rauschen mit mehreren Aktionspotentialen antwortet, wird mit zunehmender Rauschamplitude größer. Durch den Schwellenmechanismus der nachgeschalteten Richtungsdetektoren wird das Rauschen unterdrückt. Die Wahrscheinlichkeit, daß ein Richtungsdetektor auf reines Rauschen mit der Generierung von Aktionspotentialen reagiert, ist im Vergleich zu den retinalen Ganglienzellen sehr niedrig. Dennoch wirkt sich das Rauschen auch bei den Richtungsdetektoren nachteilig auf das Antwortverhalten aus. Die Wahrscheinlichkeit, daß ein Richtungsdetektor kein Aktionspotential generiert, obwohl sich eine bewegte Kante in seinem rezeptiven Feld befindet, wächst mit steigender Rauschamplitude. Dadurch wird eine Entscheidung, ob ein Signal oder nur Rauschen vorliegt, erschwert.

In den Verteilungen der Abbildung 3.48 ist zusätzlich noch eine Schwelle $\Theta_{95\%}$ dargestellt. Diese Schwelle gibt an, ab welcher Aktionspotentialanzahl pro Aktionspotentialfolge mit einer Wahrscheinlichkeit von mindestens 95% ein Signal vorliegt. Bei den Verteilungen der relativen Häufigkeit retinaler Ganglienzellen wird diese Schwelle mit zunehmender Rauschamplitude höher. Reichen bei einer Rauschamplitude von $V^{\text{noise}} = 20$ bereits zwei Aktionspotentiale aus, um mit einer Sicherheit von mindestens 95% ein Signal detektiert zu haben, so sind bei $V^{\text{noise}} = 90$ mindestens sechs Aktionspotentiale pro Aktionspotentialfolge dafür nötig. Bei den Verteilungen der Richtungsdetektoren, sowohl mit als auch ohne laterale Linking-Verbindungen, ist die Schwelle $\Theta_{95\%}$ sehr niedrig. Bereits bei zwei Aktionspotentialen pro Aktionspotentialfolge ist mit einer Sicherheit von 95% das Signal die Ursache. Offen bleibt dabei die Frage, mit welcher Wahrscheinlichkeit p_\circ die Modellneuronen in der Lage sind, die Kante zu detektieren. Hierbei zeigt sich ein weiterer Vorteil der lateralen Linking-Verbindungen.

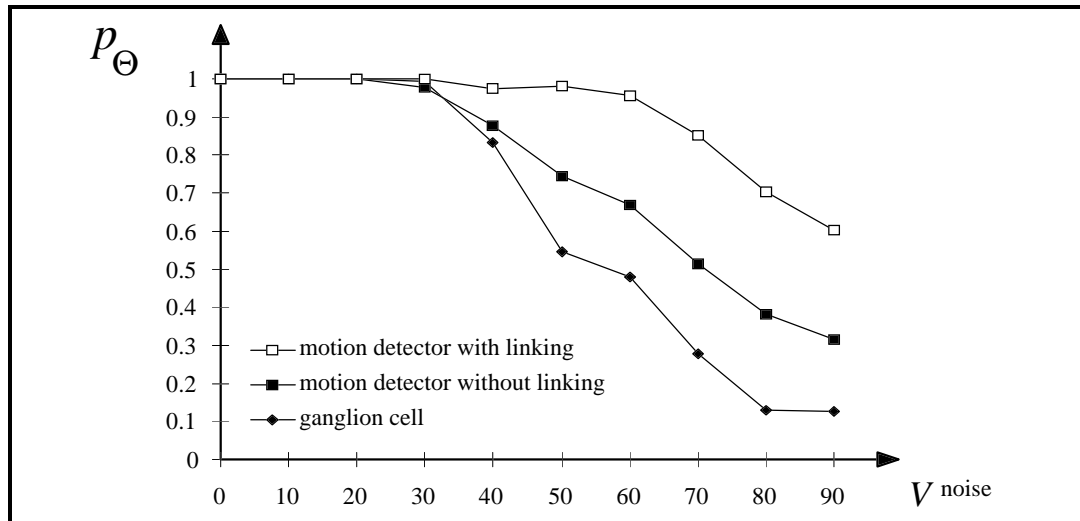


Abb. 3.49: Detektionsfähigkeit von Ganglienzellen und Richtungsdetektoren. Aufgetragen sind die Wahrscheinlichkeiten p_{Θ} , die angeben, wie groß der Anteil der 0° -Richtungsdetektoren ist, deren Aktionspotentialanzahl pro Aktionspotentialfolge so groß ist, daß mit 95-prozentiger Sicherheit eine Kante detektiert wurde, in Abhängigkeit zur Rauschamplitude V^{noise} .

Die in der Abbildung 3.49 aufgetragenen Wahrscheinlichkeiten p_{Θ} geben an, wie groß der Anteil der 0° -Richtungsdetektoren ist, deren Aktionspotentialanzahl pro Aktionspotentialfolge so groß ist, daß mit 95-prozentiger Sicherheit eine Kante detektiert wurde. Die retinalen Ganglienzellen zeigen hier die schlechteste Detektionsfähigkeit. Bereits ab einer Rauschamplitude von $V^{\text{noise}} = 50$ erfüllen nur noch die Hälfte der Ganglienzellen dieses Kriterium. Besser schneiden die Richtungsdetektoren, insbesondere mit lateraler Linking-Kopplung, ab. Bis zu einer Rauschamplitude von $V^{\text{noise}} = 60$ sind noch weit über 90% der lateral gekoppelten Richtungsdetektoren in der Lage, ausreichend viele Aktionspotentiale zu generieren. Bei der maximalen Rauschamplitude von $V^{\text{noise}} = 90$ sind es mit lateraler Kopplung immer noch doppelt so viele Modellneuronen wie ohne laterale Kopplung. Dieses Ergebnis beruht auf der Möglichkeit der lateralen Interaktion der Modellneuronen, deren receptive Felder entlang der Kante liegen. Modellneuronen, für die dies nicht gilt, können sich nicht oder nur gering wechselseitig unterstützen. Die verbesserten Detektionseigenschaften sind somit von der Kantenlänge abhängig.

3.2.2.2 Simulationsergebnisse mit realen Szenen

Standen im letzten Kapitel noch Simulationen mit künstlichen Szenen als Eingabe im Vordergrund, so sind dies nun Simulationen mit realen Szenen. Mit dem vorgestellten neuronalen Netzwerk zur Bewegungsdetektion lassen sich nicht nur psychophysische Effekte erklären, sondern auch sich bewegende Objekte in realen Szenen, wie es in der technischen Bildverarbeitung erforderlich ist, detektieren. Dazu werden die gleichen Parameter, wie sie in den Tabellen 3.2, 3.3 und 3.4 aufgeführt sind, benutzt. Mit realen Szenen sind in diesem Zusammenhang keine konstruierten Szenen mit realen Objekten gemeint, vielmehr beinhalten die in den nachfolgenden Simulationen verwendeten realen Szenen typische Situationen des gewöhnlichen Straßenverkehrs. Bei den Aufnahmen der Szenen war die Kamera nicht statisch

positioniert, sondern bewegte sich selbst in einem Kraftfahrzeug mit. Der größte Anteil an Sequenzen wurde mit einer handelsüblichen Videokamera (Baujahr 1990) aufgenommen.

Die Aufgabe des Transientensystems in der technischen Anwendung ist die Detektion „interessanter“ Bildbereiche - das sind hier die sich bewegenden Objekte -, so daß über eine Aufmerksamkeitssteuerung (z.B. Pauly et al. 1997, Pauly et al. 1998) die Kamera so ausgerichtet werden kann, daß ein hochauflösendes Kontur/Form-System (z.B. Weitzel et al. 1996, Weitzel et al. 1997a, Weitzel et al. 1997b, Weitzel 1998) eine Objektklassifizierung vornehmen kann. Im Straßenverkehr sind solche interessanten Objekte zum Beispiel Fußgänger, Verkehrszeichen und andere Kraftfahrzeuge. Aus diesem Grund wird in den nachfolgenden Simulationen die Detektierbarkeit dieser Objekte im Straßenverkehr, sofern sie sich relativ zur Kamera bewegen, untersucht.

Kraftfahrzeuge

Zunächst soll anhand der Szene „Gerade“ die Verbesserung der Detektion sich bewegender Objekte mittels lateraler Kopplungen über Linking-Verbindungen zwischen den Richtungsdetektoren dargestellt werden. Weiterhin wird der Vorteil einer zusätzlichen Kopplung zwischen ON- und OFF-Pfad gezeigt. Die Bilder sind 288 Pixel breit und 183 Pixel hoch. Da es sich bei den Bildern um sogenannte Halbbilder handelt, wurden jeweils zwei Durchläufe, zum einen mit den geradzahligen und zum anderen mit den ungeradzahligen Bildern, durchgeführt. Damit ergibt sich eine Simulationsdauer von $t = 1472$ Bin bzw. $t = 1504$ Bin. Wegen der beschränkten Anzahl an Modellneuronen wurde ein Abtastraster mit 60 Verarbeitungseinheiten pro Zeile und 40 Verarbeitungseinheiten pro Spalte in der mittleren und der feinen Auflösung benutzt. Daraus ergibt sich eine Modellneuronenanzahl von 124 800 und eine Verbindungsanzahl von unter 1 000 000.

Die kumulierten Aktivitäten der Richtungsdetektoren der mittleren Auflösung sind in der Abbildung 3.50 aufgetragen. Die Ergebnisse der Simulation mit den ungeradzahligen Eingabebildern befinden sich in Abbildung 3.50 A und B, die der geradzahligen in Abbildung 3.50 C und D. Für den ON- (Abb. 3.50 A und C) und den OFF- (Abb. 3.50 B und D) Pfad sind die kumulierten Aktivitäten getrennt berechnet worden. Die Simulationsergebnisse sind sehr ähnlich. Lediglich die 240°-ON-Richtungsdetektoren sind stärker aktiv als die Richtungsdetektoren des OFF-Pfades mit gleicher Richtungspräferenz. Dies resultiert aus der Bewegung des Mittelstreifens in der Eingabesequenz (siehe auch Abbildung 3.53). Der helle, vom dunklen Hintergrund umgebene Mittelstreifen bewegt sich im Bild nach links unten. Dadurch aktiviert er besonders die 240°-ON-Richtungsdetektoren. Am aktivsten sind aber die 0°-Richtungsdetektoren. Sie detektieren die Bewegung des Kraftfahrzeuges in die Bildmitte. Die 3-dimensionale Bewegung des Kraftfahrzeuges in die Tiefe ist auf dieser Verarbeitungsebene nicht extrahierbar.

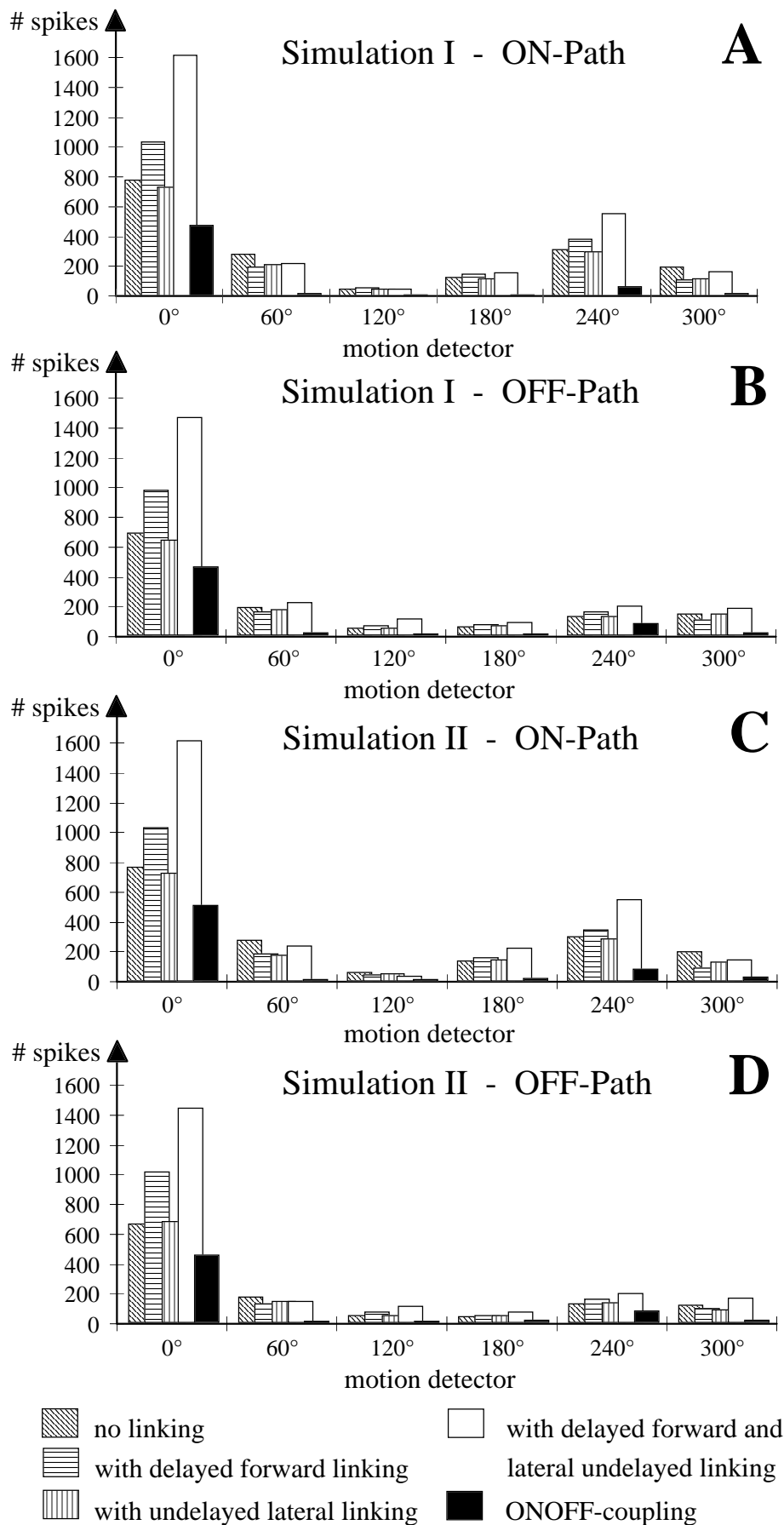


Abb. 3.50: Kumulierte Aktivitäten der Richtungsdetektoren (Sequenz „Gerade“). Simulation I (A, B) wurde mit den ungeradzahligen, Simulation II (C, D) mit den geradzahligen Bildern der Sequenz „Gerade“ durchgeführt.

Ohne laterale Kopplungen zwischen den Richtungsdetektoren sind zwar die 0° -Detektoren am aktivsten, doch die anderen sind ebenfalls noch aktiv. Dies wird auch in der Abbildung 3.51 A deutlich, in der die Aktivitäten der OFF-Richtungsdetektoren in vier aufeinanderfolgenden Bildern als Länge der Pfeile dargestellt sind (Kopecz et al. 1996). Das Hinzufügen von unverzögerten lateralen Linking-Kopplungen alleine führt zunächst zu keiner signifikanten Verbesserung des Detektionsergebnisses (Abbildungen 3.50 und 3.51 B). Die verzögerten Linking-Kopplungen führen hingegen schon zu einer Erhöhung der 0° -Richtungsdetektoraktivität um etwa 30%, ohne dabei die Aktivität der anderen Richtungsdetektoren zu erhöhen (Abbildungen 3.50 und 3.52 A). Ausgenommen sind hiervon die 240° -Richtungsdetektoren, da die kontinuierliche Bewegung des Mittelstreifens eine Erhöhung der Aktivität dieser Detektoren über die Voraktivierung ebenso ermöglicht. Die Kombination beider Linking-Kopplungen führt indessen zu einer etwa um den Faktor 2 stärkeren Erhöhung der 0° -Richtungsdetektoraktivität (Abbildungen 3.50, 3.52 B und 3.53 A). Zur Unterdrückung der schwach aktiven Richtungsdetektoren dient die nachfolgende Kopplung der beiden ON- und OFF-Verarbeitungspfade. Wie in der Abbildung 3.53 B zu sehen ist, sind im Bereich des überholenden Kraftfahrzeuges fast ausschließlich 0° -ONOFF-Richtungsdetektoren aktiv. Ferner wird der vordere Teil des Mittelstreifens von den 240° -ONOFF-Richtungsdetektoren detektiert. Der hintere Teil des Mittelstreifens bewegt sich mit ungefähr $v = 2$ Pixel/Bild im Gegensatz zur Geschwindigkeit des Vorderteils ($v = 5$ Pixel/Bild) für die Richtungsdetektoren dieser Auflösung zu langsam, um zu einer ausreichend starken Aktivität zu führen.

Mit zunehmender Entfernung des überholenden Fahrzeugs von der Kamera wird dessen projizierte Geschwindigkeit in das Eingabebild geringer, wodurch die Aktivität der Richtungsdetektoren der mittleren Auflösung zurückgeht. Für die Richtungsdetektoren der feinen Auflösung reicht diese Geschwindigkeit allerdings noch aus, wie die Abbildung 3.54 zeigt. Auch in dieser Abbildung wird eine Verbesserung des Detektionsergebnisses mit ONOFF-Interaktion bezüglich der Eindeutigkeit der Bewegungsbestimmung deutlich.

Zur Vermeidung von Gefahren im Straßenverkehr durch überholende Fahrzeuge ist aber zunächst die grobe Auflösung nützlicher als die feine oder mittlere, da sich mit dieser das überholende Fahrzeug früher detektieren läßt. Eine rechtzeitige Detektion könnte dann zu seiner Verfolgung und zur Schätzung seiner weiteren Fahrlinie genutzt werden, um zum Beispiel ein zu knappes Einscheren des Überholenden zu erkennen oder ein zu frühes Starten des eigenen Überholvorganges zu unterbinden.

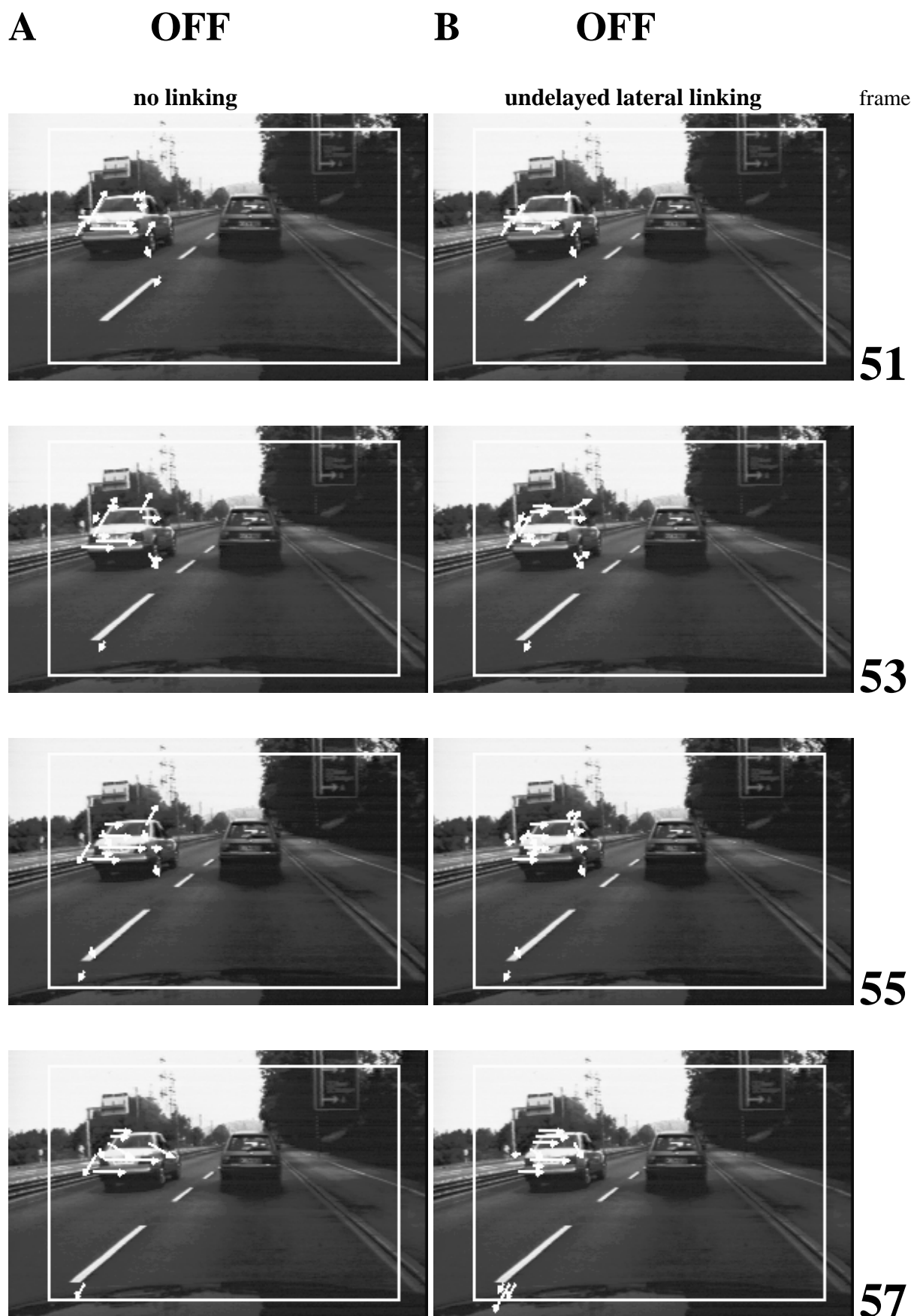


Abb. 3.51: Simulation „Gerade“ I. Dargestellt sind die Aktivitäten der OFF-Richtungsdetektoren ohne (A) und mit unverzögerten Linking-Kopplungen (B) als Pfeillängen in vier aufeinanderfolgenden Bildern. Der Verarbeitungsbereich wird durch das helle Rechteck markiert.

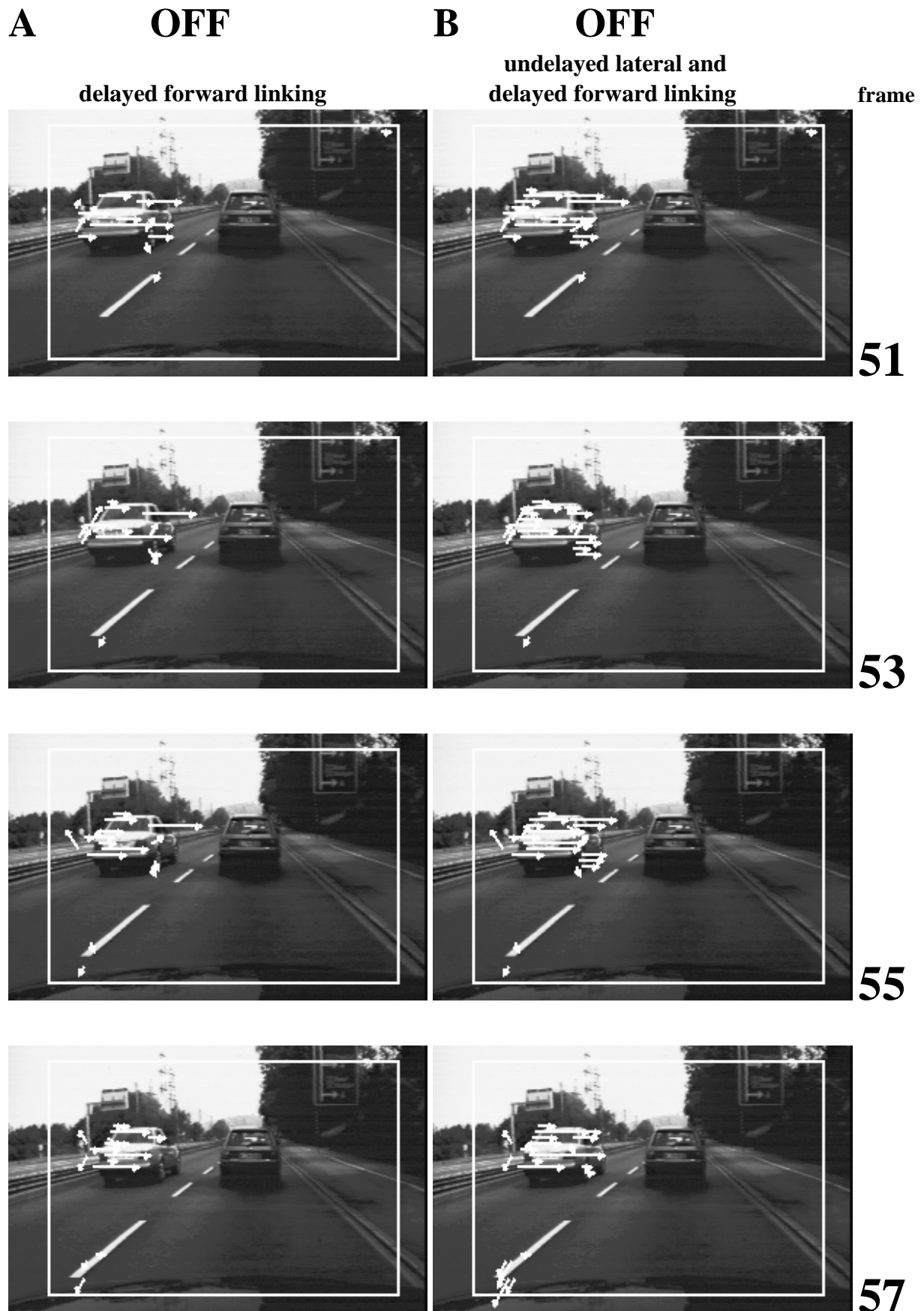


Abb. 3.52: Simulation „Gerade“ II. Dargestellt sind die Aktivitäten der OFF-Richtungsdetektoren der mittleren Auflösung mit verzögerten (A) und einer Kombination aus verzögerten und unverzögerten (B) Linking-Kopplungen als Pfeillängen in vier aufeinanderfolgenden Bildern.

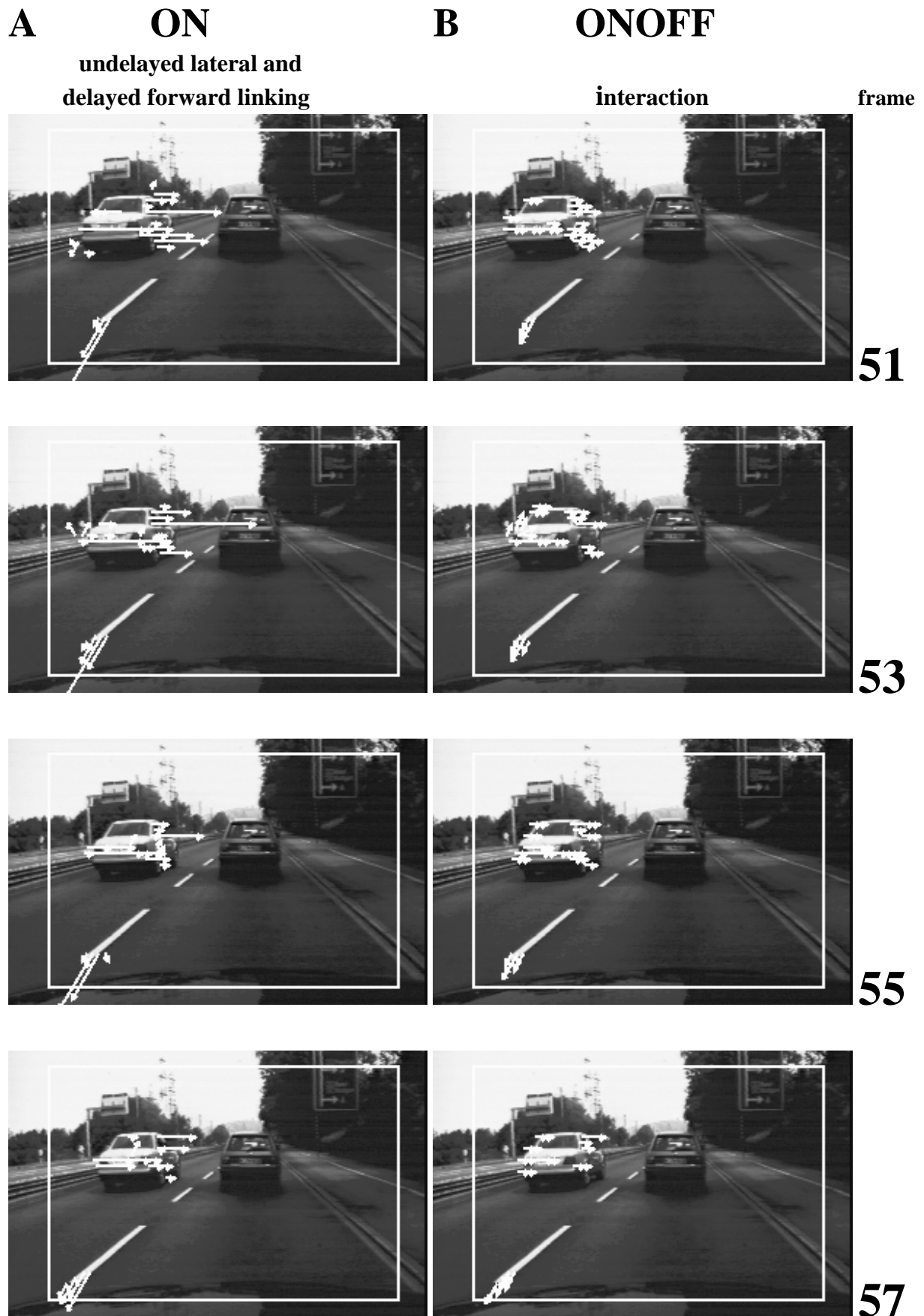


Abb. 3.53: Simulation „Gerade“ III. Dargestellt sind die Aktivitäten der ON-Richtungsdetektoren mit verzögerten und unverzögerten Linking-Kopplungen (**A**) sowie die Aktivitäten der ONOFF-Richtungsdetektoren (**B**) als Pfeillängen in vier aufeinanderfolgenden Bildern.

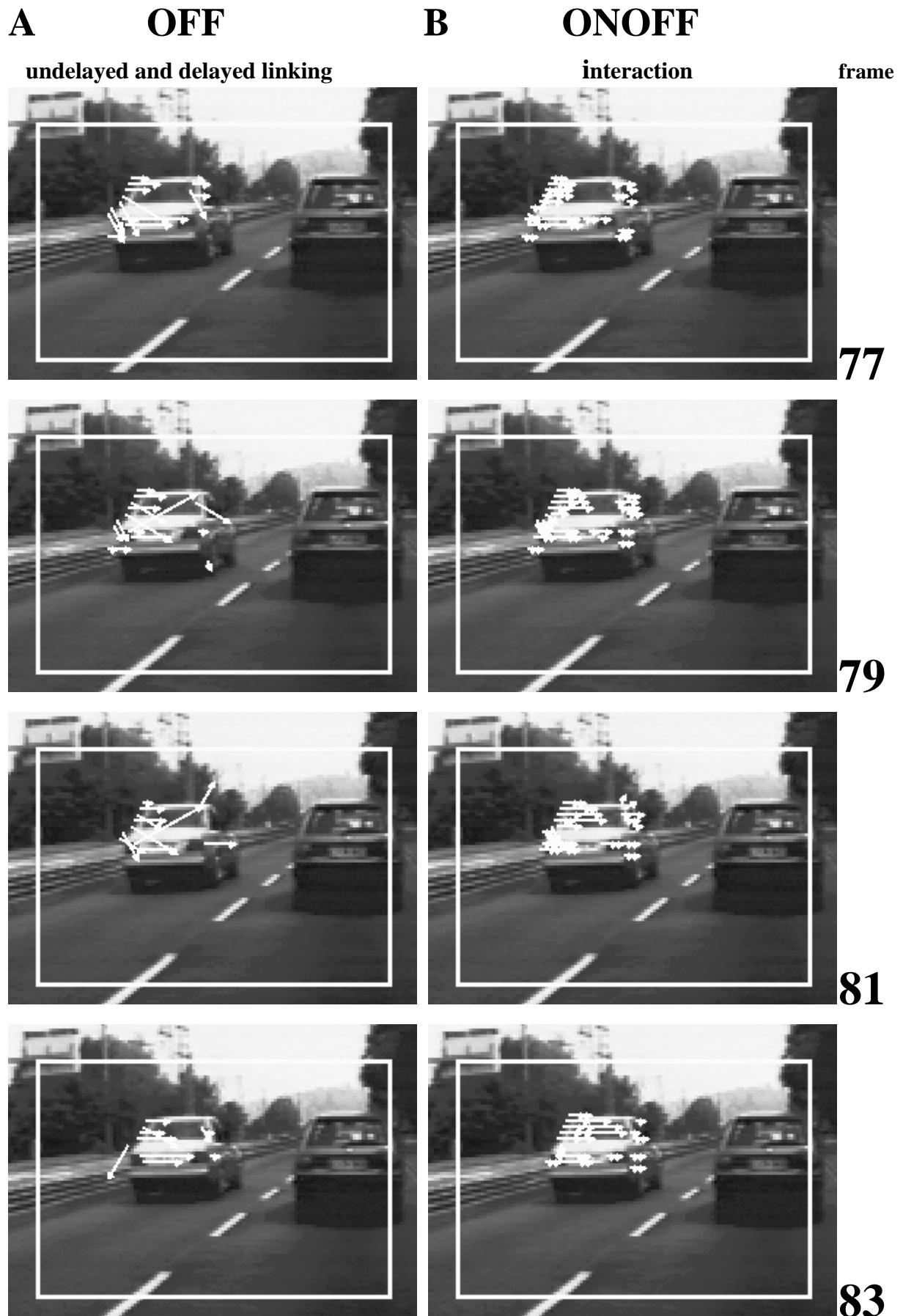


Abb. 3.54: Simulation „Gerade“ IV. Aktivitäten der OFF-Richtungsdetektoren der feinen Auflösung mit Linking-Kopplungen (**A**) sowie die Aktivitäten der ONOFF-Richtungsdetektoren (**B**). Die Abbildung wurde im Vergleich zu den vorherigen um den Faktor 2 vergrößert dargestellt.

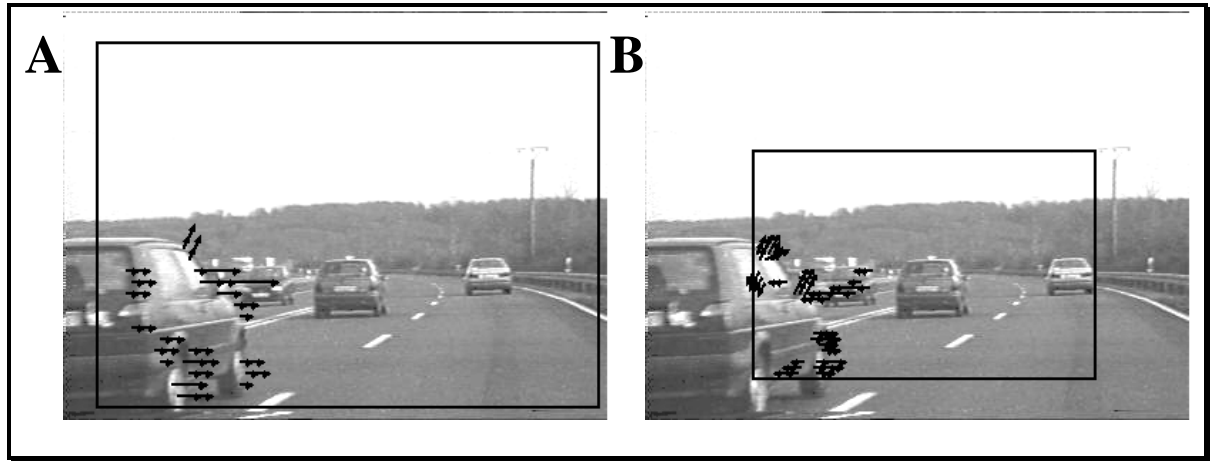


Abb. 3.55: Simulation „Überholendes Fahrzeug“. In **A** sind die ONOFF-Richtungsdetektorantworten der groben, in **B** die der mittleren Auflösung dargestellt. Der Verarbeitungsbereich wird durch das schwarze Rechteck markiert. Die räumliche Position des Verarbeitungsbereiches der mittleren Auflösung wurde in diesem Beispiel nach unten verschoben, damit ein Vergleich zwischen den Richtungsdetektorantworten der groben und mittleren Auflösung gezogen werden kann. Die Pfeillänge gibt die Stärke der Aktivierung an.

Die Detektion eines überholenden Fahrzeuges durch ONOFF-Richtungsdetektoren der groben Auflösung ($d = 8$ Pixel, 44×32 Verarbeitungseinheiten pro Schicht, Bildgröße 384 Pixel \times 288 Pixel) ist in der Abbildung 3.55 A dargestellt. Fast alle aktiven Detektoren zeigen eine Bewegungsrichtung in die Bildmitte an. Die Kamera könnte somit durch eine Aufmerksamkeitssteuerung zur genaueren Analyse auf das Fahrzeug gerichtet werden und dieses verfolgen. In der groben Auflösung wird jedoch das entgegenkommende Fahrzeug nicht detektiert. Die Geschwindigkeit dieses Objektes in den Eingabebildern ist zu gering. Die ONOFF-Richtungsdetektoren der mittleren Auflösung sind dagegen in der Lage, dieses Fahrzeug zu detektieren, wie in der Abbildung 3.55 B zu sehen ist. Die hohe Geschwindigkeit des vorbeifahrenden Fahrzeugs führt zu einer Vergrößerung des Korrespondenzproblems in der mittleren Auflösung. Die Erzeugung eindeutiger Richtungsdetektorantworten durch die Kopplung unterschiedlicher Auflösungen wird in Kapitel 3.3 untersucht.

Zur Überprüfung der Leistungsfähigkeit, überholende Kraftfahrzeuge mit den Richtungsdetektoren der groben Auflösung zu detektieren, wurden insgesamt 38 verschiedene Überholvorgänge ausgewertet. Zur Auswertung wurden jeweils zwischen zwei aufeinanderfolgenden Bildern $P(p)$ und $P(p+1)$ die Anzahl S_p bzw. s_p der Aktionspotentiale insgesamt (Gl. 3.48) und die durch das überholende Fahrzeug verursachten (Gl. 3.49) bestimmt.

$$S_p = \sum_{t=p-32 \text{ Bin}}^{(p+1) \cdot 32 \text{ Bin} - 1 \text{ Bin}} MUA_{\text{Fahrzeug}}(t) \quad (\text{Gl. 3.48})$$

$$S_p = \sum_{t=p-32 \text{ Bin}}^{(p+1) \cdot 32 \text{ Bin} - 1 \text{ Bin}} MUA_{\text{gesamt}}(t) \quad (\text{Gl. 3.49})$$

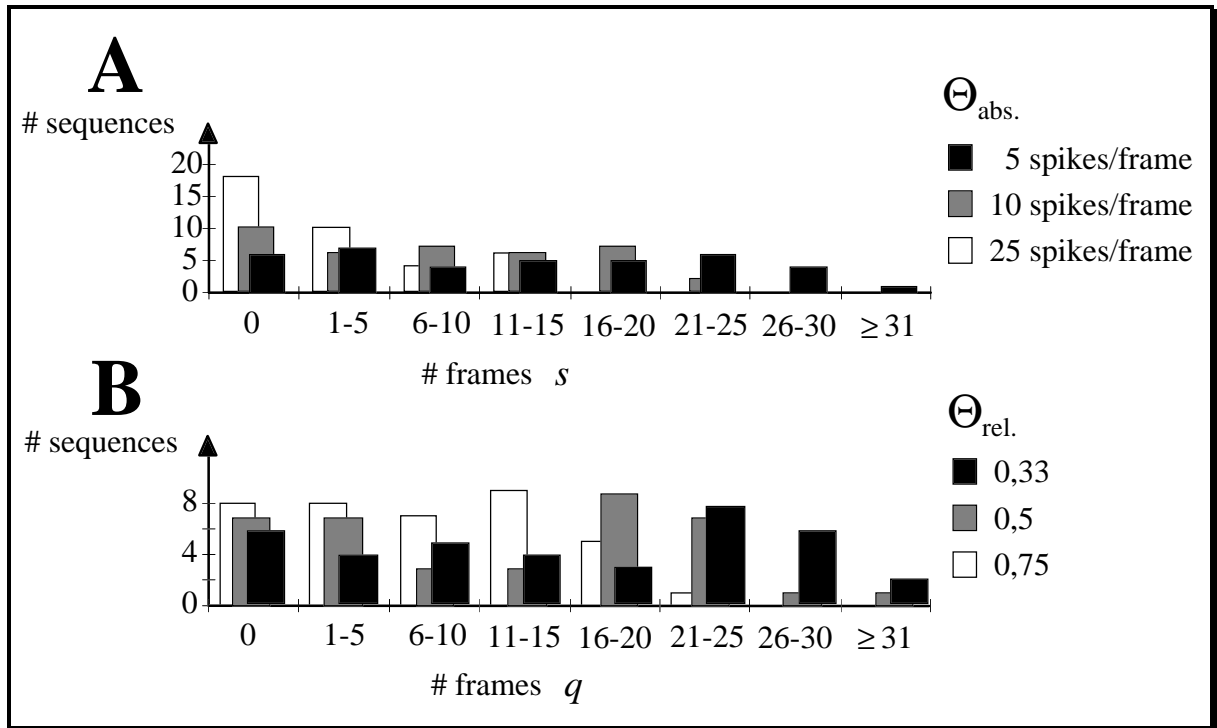


Abb. 3.56: Histogramm „Überholende Fahrzeuge“. Anzahl der Sequenzen, in denen s bzw. q Bilder eine ausreichend hohe, durch die Parameter $\Theta_{abs.}$ (A) und $\Theta_{rel.}$ (B) charakterisierte ONOFF-Richtungsdetektoraktivität generieren. Nähere Erläuterungen im Text.

In beiden Gleichungen ist $p \in N$, mit $p \cdot 32 \text{ Bin} < t_{\max}$. MUA ist im Falle des Fahrzeugs die aufsummierte Aktivität (vgl. Gl. 3.18) der 0° -ONOFF-Richtungsdetektoren, die das Fahrzeug repräsentieren, und im anderen Fall die Summe der Aktivitäten aller ONOFF-Richtungsdetektoren.

Die Auswertung ist in Abbildung 3.56 graphisch dargestellt. Die Anzahl an Bildern einer Sequenz, bei denen s_p größer oder gleich als ein festgelegter Schwellenwert $\Theta_{abs.}$ ist, wird mit s bezeichnet. Damit läßt sich die Anzahl der Sequenzen, in denen s Bilder dieses Schwellenkriterium erfüllen, in Form eines Histogrammes auftragen. So gibt es zum Beispiel fünf Sequenzen, in denen 11 bis 15 Bilder eine Mindestanzahl von fünf Aktionspotentialen pro Bild verursachen.

Die absolute Anzahl an Aktionspotentialen pro Bildwechsel - abgesehen von einer Mindestanzahl an generierten Aktionspotentialen - sagt aber noch nichts darüber aus, inwieweit eine Zielauswahl durch eine Aufmerksamkeitssteuerung erfolgen kann. Soll die Zielauswahl ausschließlich aufgrund der Aktivität der Richtungsdetektoren erfolgen, so muß die vom Zielobjekt verursachte Aktivität s_p im Verhältnis zur Gesamtaktivität S_p betrachtet werden. Mit diesem daraus resultierenden Quotienten q_p wird dann ein weiterer Schwellenvergleich mit $\Theta_{rel.}$ vorgenommen. Das Ergebnis der Auswertung ist in Abbildung 3.56 B als Histogramm dargestellt. Hierbei ist q die Anzahl der Bilder, bei denen $q_p \geq \Theta_{rel.}$ ist. Es zeigt sich, daß in 63% der Sequenzen es mindestens sechs Bilder gibt, in denen die durch das Zielobjekt verursachte Aktivität mindestens die Hälfte der Gesamtaktivität darstellt ($\Theta_{rel.} = 0.5$). Für das schärfere Kriterium ($\Theta_{rel.} = 0.75$) sind es immerhin noch 58% und für das schwächere Kriterium ($\Theta_{rel.} = 0.33$) sogar 74% der Sequenzen. Ein Großteil der überholenden Fahrzeuge wird somit vom Netzwerk detektiert, so daß eine Zielauswahl aufgrund der relativen Aktivität

stattfinden kann. Die von den Richtungsdetektoren der groben Auflösung nicht detektierbaren Zielobjekte sind in der Regel zu langsam überholende Fahrzeuge. Aufgrund der geringen Relativbewegung zum eigenen Fahrzeug beinhalten diese Überholmanöver allerdings nur ein geringes Gefahrenpotential.

Verkehrszeichen

Im Straßenverkehr stellen Verkehrszeichen in Form von Schildern oder Markierungen auf der Fahrbahn für die Auswertung durch ein Fahrer-Assistenz-Sehsystem eine weitere Gruppe interessanter Objekte dar. Ein aktives Sehsystem könnte nach „Erregung der Aufmerksamkeit“ auf ein Verkehrszeichen schwenken, es verfolgen und es somit nahezu statisch im hochauflösenden Verarbeitungsbereich eines Kontur/Form-Systems halten. Ein Kontur/Form-System hat dann die Möglichkeit, das Objekt genau zu analysieren und zu klassifizieren. Die Abbildung 3.57 zeigt die Detektion eines großen Verkehrsschildes an der linken Straßenseite und einer einfachen Markierung auf der Mitte der Fahrbahn durch Richtungsdetektoren der mittleren Auflösung in sechs aufeinanderfolgenden Bildern. Da es keine 270°-Richtungsdetektoren gibt, wird die Fahrbahnmarkierung durch 240°- und 300°-ONOFF-Richtungsdetektoren detektiert. Aufgrund der zu niedrigen Kameraposition könnte die Fahrbahnmarkierung jedoch nur kurz von dem aktiven Kamerasystem verfolgt werden. Im Bereich des Schildes sind hauptsächlich Detektoren mit einer Richtungspräferenz von 180°, aber auch welche mit Richtungspräferenzen von 60° und 240° aktiv. Diese Detektoraktivitäten resultieren aus den größer werdenden Projektionen des Schildes in den Eingabebildern aufgrund der Verringerung des Abstandes zwischen Kamera und Schild. Modellneuronen mit einer Richtungspräferenz von 180° detektieren bevorzugt den linken Teil des Schildes, wohingegen die 60°- und 240°-Richtungsdetektoren eher durch den oberen bzw. den unteren Teil des Schildes aktiviert werden. Bei Kenntnis der Zusammengehörigkeit der neuronalen Aktivität, wie sie durch ein Kontur/Form-System herausgearbeitet werden könnte, ist somit aus der 2-dimensionalen Projektion auch eine Tiefeninformation, in diesem Fall das relative Näherkommen des Schildes, zu gewinnen.

In den nachfolgenden Simulationen soll gezeigt werden, daß bereits die Relativbewegung der Schilder zur sich bewegenden Kamera ausreicht, um diesen Bildbereich durch eine erhöhte Aktivität der Richtungsdetektoren für eine Aufmerksamkeitssteuerung zu markieren. Dazu wurden insgesamt 60 willkürlich auf einer Autobahn (mit einer Geschwindigkeit von etwa 100 km/h) aufgenommene Sequenzen benutzt. Die Abbildung 3.58 zeigt eines von 60 zu detektierenden Schildern. Wie in der Abbildung deutlich wird, werden nicht nur Schilder, sondern auch Leitpfosten von den Richtungsdetektoren detektiert. Die Bewegungsrichtung anderer Objekte, wie zum Beispiel die einzelner Büsche, können ebenfalls vom Transientensystem detektiert werden. Die Entscheidung, ob es sich um einen Busch oder um ein Verkehrsschild handelt, ist dann nach einem Kameraschwenk auf das Objekt vom Kontur/Form-System zu treffen.

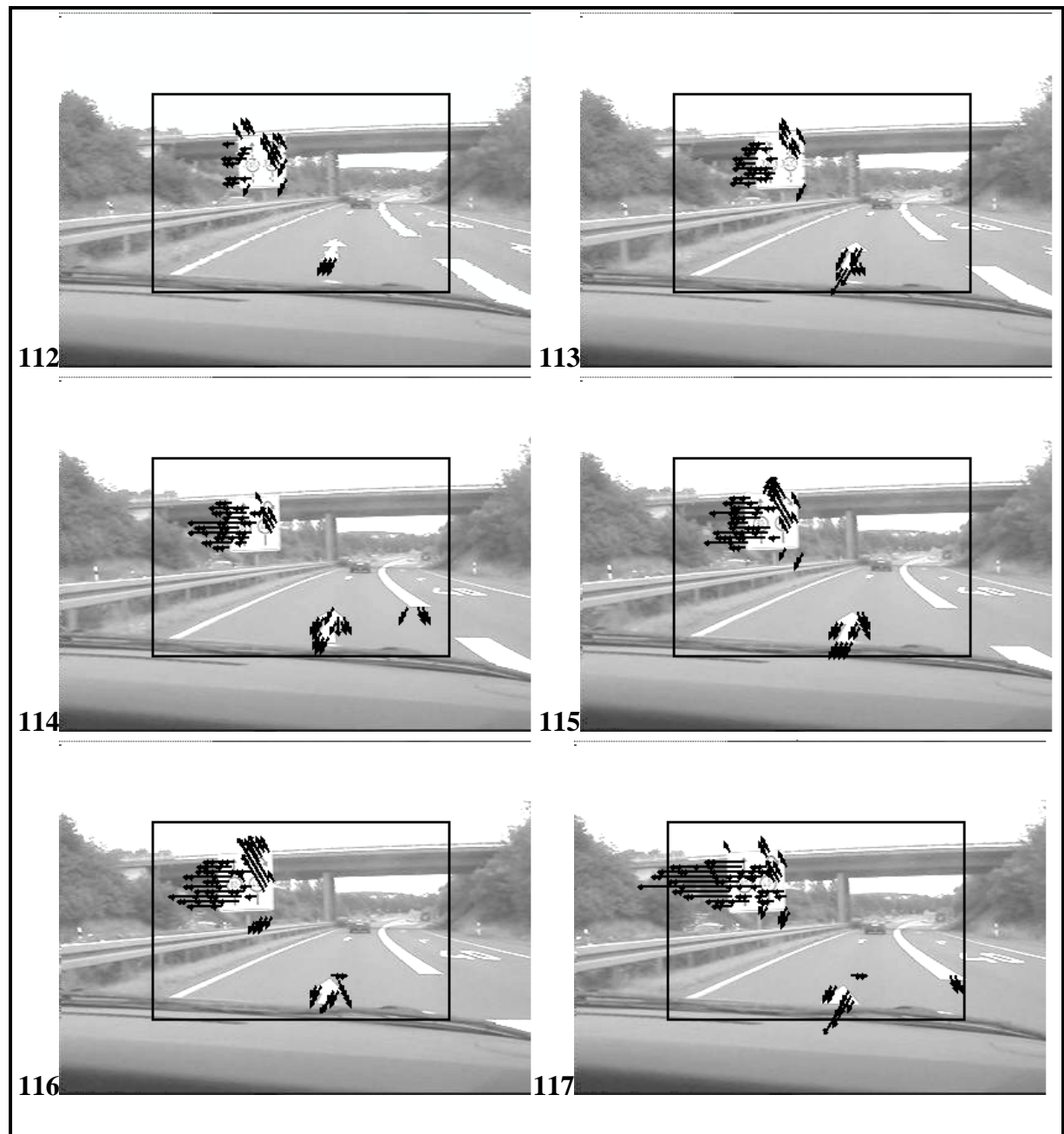


Abb. 3.57: Detektion von Verkehrszeichen. Die Pfeile stellen die Aktivität der ONOFF-Richtungsdetektoren der mittleren Auflösung dar. Die Detektion des Schildes durch 60°-, 180°- und 240°-ONOFF-Richtungsdetektoren weist bei Kenntnis der Zusammengehörigkeit auf ein größer werdendes Objekt hin.

Entsprechend der Auswertung der Überholvorgänge werden auch aus den Ergebnissen der Simulationen mit den Verkehrsschildern s und q bestimmt. Die Auswertung ist als Histogramm in Abbildung 3.59 aufgetragen. Wird eine absolute Schwelle von $\Theta_{\text{abs.}} = 5$ Aktionspotentiale/Bild zugrunde gelegt, dann werden nur 5% der Verkehrsschilder von den ONOFF-Richtungsdetektoren nicht detektiert. 80% der Verkehrsschilder erzeugen in mindestens einem Bild eine Mindestaktivität von 50% der Gesamtaktivität. Damit kann die Richtungsdetektion auch für Verkehrsschilder als Eingabe für eine Aufmerksamkeitssteuerung dienen.



Abb.3.58: Detektion eines Schildes und eines Leitpfostens. Das im Bild dargestellte Verkehrsschild ist eines von 60, die zur Untersuchung der Detektierbarkeit am rechten Straßenrand stehender Schilder durch das Transientensystem benutzt wurden. Bei 17 Schildern befand sich ein Seitenstreifen zwischen Fahrbahn und Schild. Das schwarze Rechteck markiert den Verarbeitungsbereich. Neben der Relativbewegung des Verkehrsschildes wird auch die des Leitpfostens von den ONOFF-Richtungsdetektoren extrahiert.

Eine genauere Analyse der Simulationen, wie sie in Abbildung 3.60 dargestellt ist, zeigt eine starke Abhängigkeit des Detektionsergebnisses vom Hintergrund der Schilder. Das beste Detektionsergebnis kann erzielt werden, wenn die Schilder den Himmel als Hintergrund haben (Abb. 3.60 A). Der Hintergrund der drei nicht detektierten Schilder bestand vollständig aus Büschen oder Gras, was zu einer Reduzierung des Intensitätskontrastes führte. Die relative Bewegungsrichtung der Schilder, die nur einen geringen intrinsischen Intensitätskontrast aufweisen, ist dann von den Detektoren schwierig zu bestimmen.

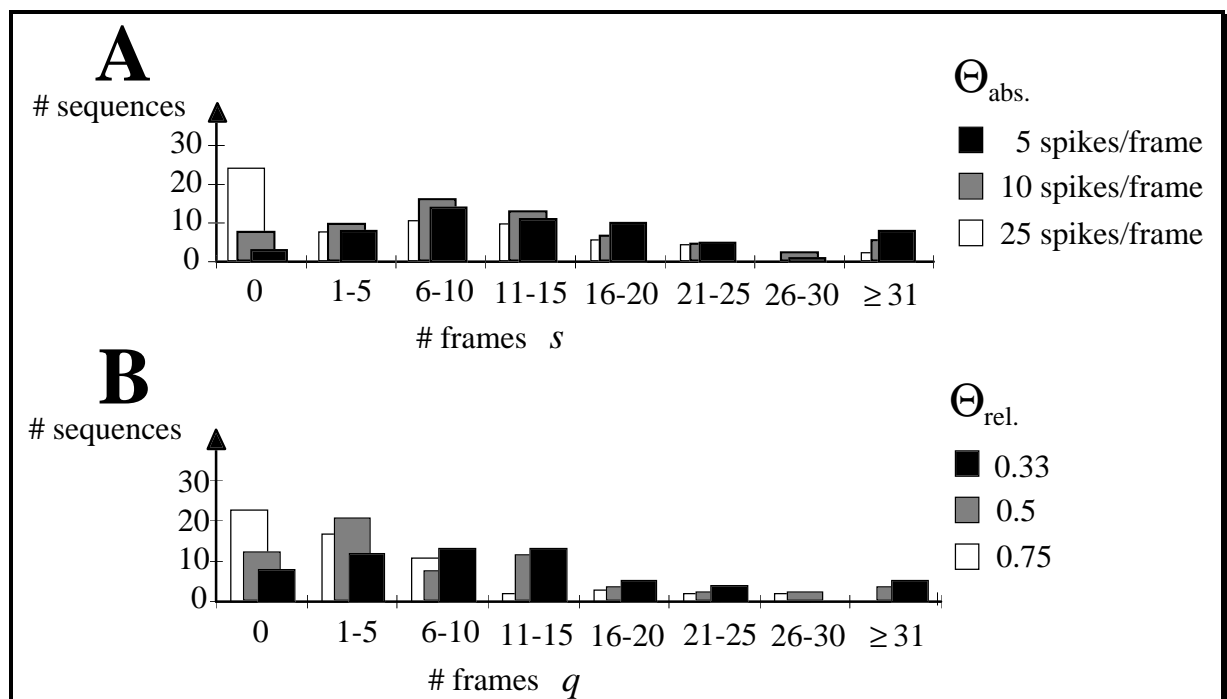


Abb. 3.59: Histogramm „Verkehrsschilder“. Anzahl der Sequenzen, in denen s bzw. q Bilder eine ausreichend hohe, durch die Parameter $\Theta_{abs.}$ (A) und $\Theta_{rel.}$ (B) charakterisierte ONOFF-Richtungsdetektoraktivität generieren. Nähere Erläuterungen im Text.

Auf das Detektionsergebnis wirkt sich auch eine geringe Größe nachteilig aus (Abbildung 3.60 B). Je größer das Verkehrsschild ist, desto höher ist die Wahrscheinlichkeit, daß seine Bewegungsrichtung von den Detektoren bestimmt werden kann. Bundesweherschilder, die Auskunft über die militärische Lastenklasse geben, stellen aufgrund ihrer geringen Größe ein Problem für das Bewegungssystem dar. Durch die geringe Größe werden nur sehr wenige Richtungsdetektoren aktiv, was zu einer schwachen Kopplung über laterale Linking-Verbindungen und somit zu einer geringen Gesamtaktivität im Bildbereich des Schildes führt. Die Abbildung 3.61 zeigt ein solches Bundesweherschild mit Büschen als Hintergrund. Ohne die Bundesweherschilder führen in 88% der Sequenzen mindestens sechs Bilder zu einer Mindestaktivität von 5 Aktionspotentialen pro Bild.

Weiterhin stellt sich der Standstreifen zwischen Fahrbahn und Schild als Nachteil für die Bewegungsrichtungsdetektion heraus. Alle drei nicht detektierten Schilder standen nicht unmittelbar neben der Straße. Dadurch werden sowohl ihre Relativgeschwindigkeit zur Kamera als auch ihre Größe verringert.

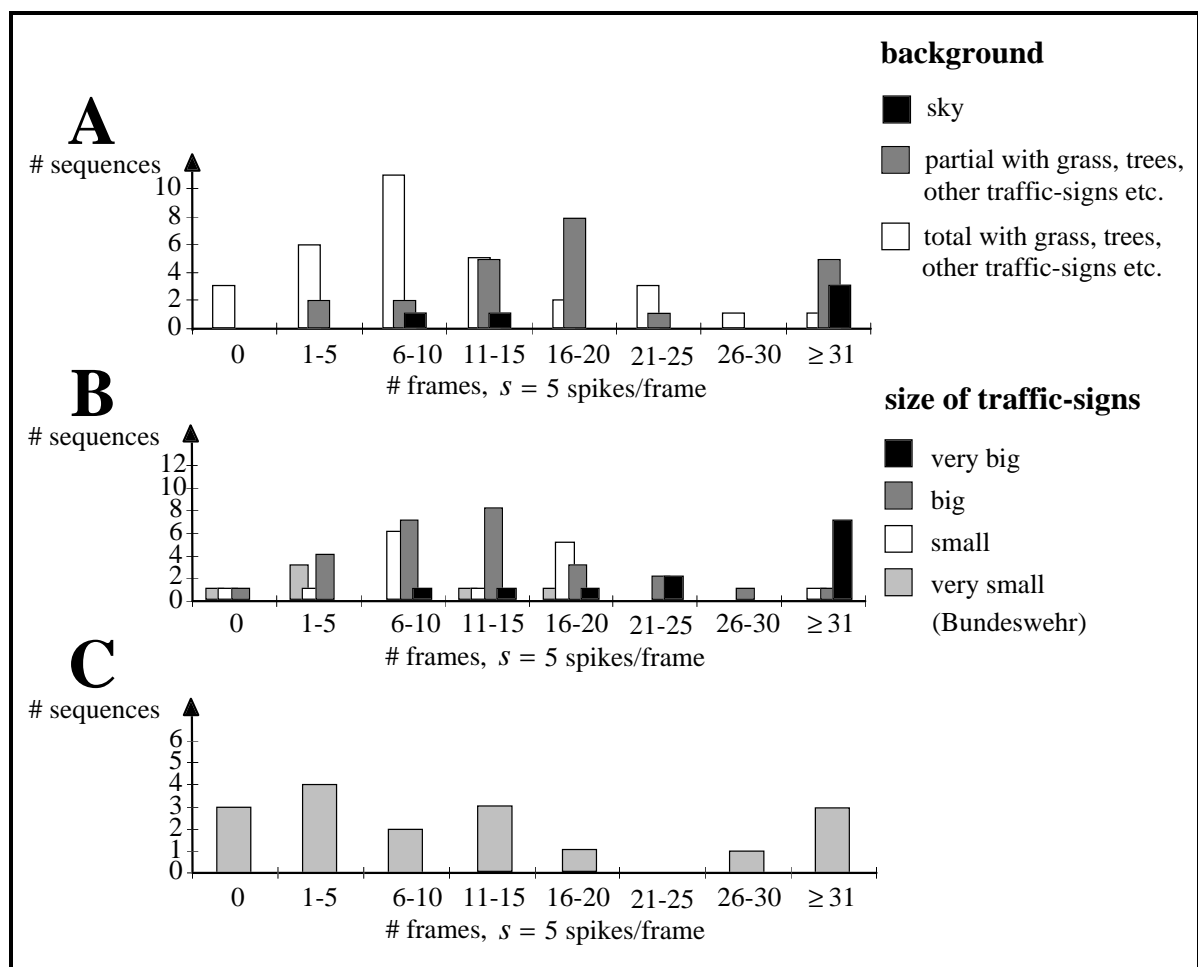


Abb. 3.60: Analyse „Verkehrsschilder“. Dargestellt ist die Abhängigkeit der ONOFF-Richtungsdetektoraktivität zum Hintergrund (A), zur Größe des Verkehrsschildes (B) und zum Vorhandensein eines zusätzlichen Standstreifens zwischen Fahrbahn und Schild (C).

Abb. 3.61: Nicht detektierbares Verkehrsschild.

In der Abbildung ist ein von den Richtungsdetektoren nicht detektiertes Verkehrsschild enthalten. Es ist ein Bundeswehrschild und befindet sich oberhalb des Leitpfostens. Aufgrund des sehr geringen Intensitätskontrastes und der geringen Größe des Schildes sind die Detektoren nicht in der Lage, dessen Bewegungsrichtung zu bestimmen.

**Fußgänger**

Aber nicht nur andere Fahrzeuge und Verkehrsschilder, sondern auch Fußgänger, insbesondere, wenn sie sich auf die Fahrbahn bewegen, stellen für ein Fahrer-Assistenz-System interessante Objekte dar. Deswegen soll in den beiden nachfolgenden Simulationen gezeigt werden, daß das neuronale Bewegungssystem auch mit dieser Problemstellung brauchbare Ergebnisse liefert. In den Abbildungen 3.62 und 3.63 werden Ausschnitte der Simulationsergebnisse gezeigt.

Die Abbildung 3.62 zeigt, daß die Bewegungsrichtung des Fußgängers detektiert wird. Während sich der Oberkörper des Fußgängers mit nahezu gleichbleibender Geschwindigkeit bewegt, verändert sich die Geschwindigkeit der Beine ständig. Dies wird auch im Simulationsergebnis deutlich. Die Bewegung des Oberkörpers wird ständig von den Detektoren extrahiert. Dies ist bei den Beinen anders. Zunächst ist das rechte Bein das Standbein und das linke wird bewegt. Die Richtungsdetektoren werden folglich nur durch das linke Bein aktiviert. Danach wird das linke Bein zum Standbein und das rechte wird bewegt, wodurch dessen Bewegungsrichtung von den Detektoren bestimmt werden kann.

Eine weitere Sequenz mit einem sich bewegenden Fußgänger ist ausschnittsweise in der Abbildung 3.63 zu sehen. In diesem Beispiel bewegt sich die Kamera nicht auf den Fußgänger zu, sondern wird langsam in dessen entgegengesetzten Richtung an ihm vorbeibewegt. Auch in diesem Beispiel können die ONOFF-Richtungsdetektoren - aufgrund der hohen Relativbewegung von etwa $v = 7$ Pixel/Bild speziell die Detektoren der groben Auflösung - die Bewegung extrahieren.

Die Simulationen dieses Kapitels haben gezeigt, daß das neuronale Netz zur Bewegungsdetektion nicht nur zur Erklärung psychophysischer Effekte herangezogen, sondern daß es auch in der technischen Bildverarbeitung - bei Beibehaltung aller Parameter mit Ausnahme der Netzgröße - erfolgreich eingesetzt werden kann. Das neuronale Netz ist somit flexibel einsetzbar.

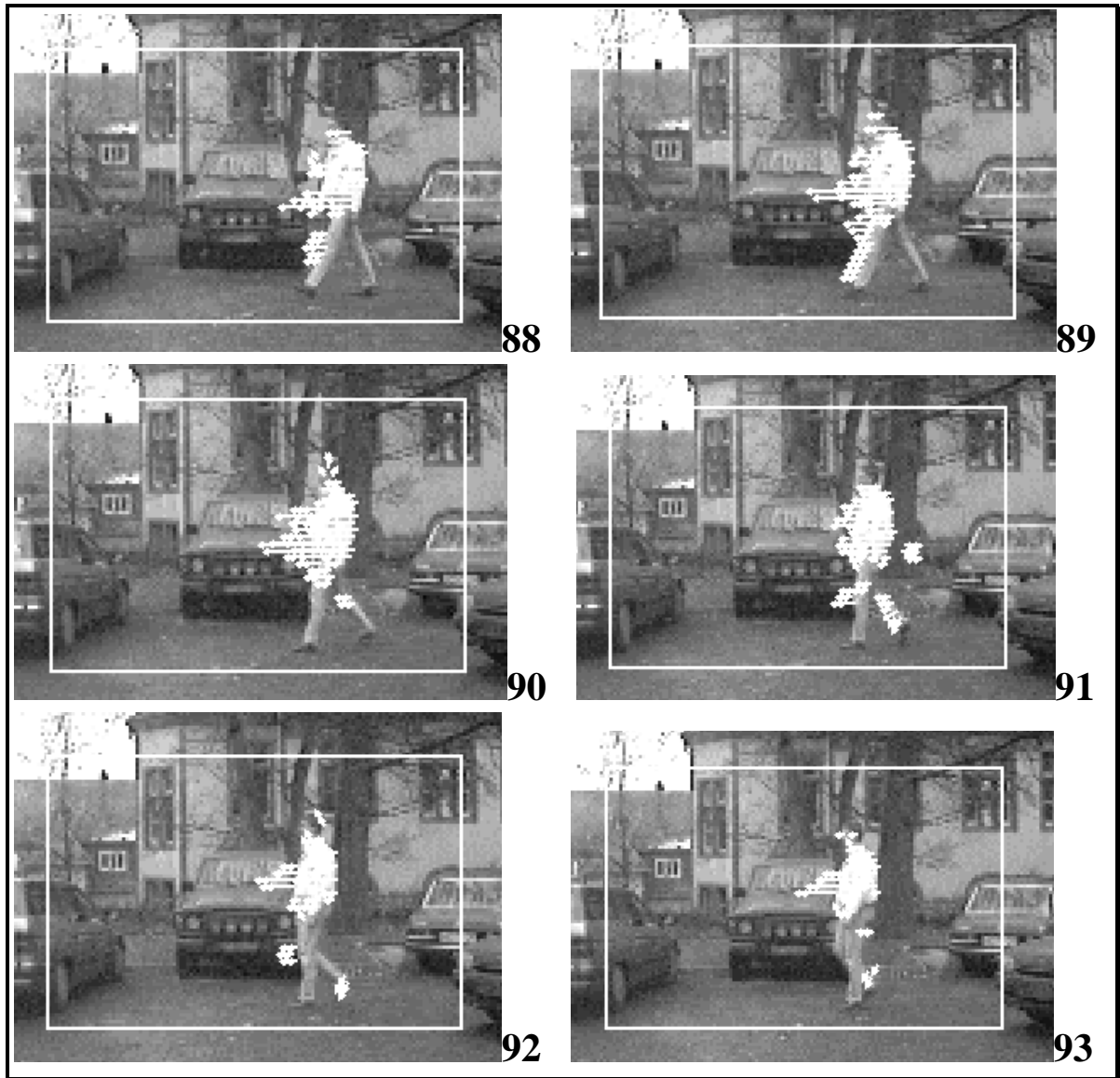


Abb. 3.62: Simulation „Fußgänger I“. Detektion eines sich bewegenden Fußgängers durch die ONOFF-Richtungsdetektoren der feinen Auflösung in sechs aufeinanderfolgenden Bildern. Der Fußgänger bewegt sich von rechts nach links durch das Bild. Die Kamera bewegt sich mit Schrittgeschwindigkeit auf den Fußgänger zu. Der Verarbeitungsbereich wird durch ein helles Rechteck markiert. Es wurden 60 x 40 Verarbeitungseinheiten der feinen Auflösung pro Schicht zur Richtungsextraktion verwendet.

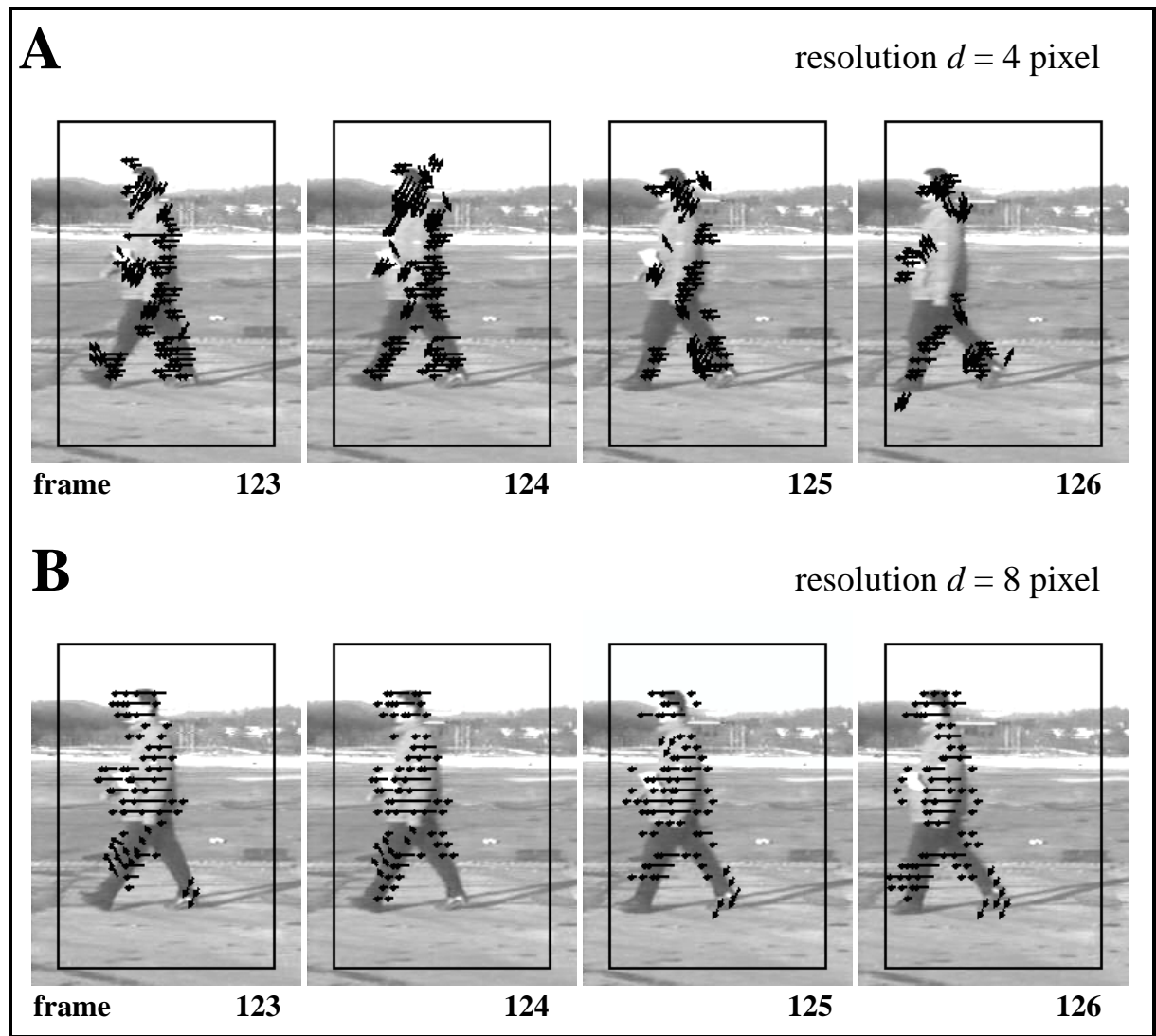


Abb. 3.63: Simulation „Fußgänger II“. Detektion eines sich von rechts nach links bewegenden Fußgängers durch ONOFF-Richtungsdetektoren der mittleren (**A**) und der groben (**B**) Auflösung. Die Kamera befindet sich in einem Fahrzeug, das sich langsam in die entgegengesetzte Bewegungsrichtung des Fußgängers bewegt. Der Verarbeitungsbereich wird durch ein schwarzes Rechteck markiert.

3.3 Kopplung von Richtungsdetektoren unterschiedlicher Auflösung

3.3.1 Modell zur Kopplung von Richtungsdetektoren unterschiedlicher Auflösung

Wie bereits in dem vorangegangenen Kapitel 3.2.2 dargestellt wurde, spiegelt sich das Korrespondenzproblem nicht nur in der mehrdeutigen Antwort der Richtungsdetektoren einer Auflösung, sondern auch in den Detektorantworten verschiedener Auflösungen wider. Diese Mehrdeutigkeiten treten bei außerhalb des für den jeweiligen Richtungsdetektor einer bestimmten Auflösung optimalen Geschwindigkeitsbereiches liegender Objektgeschwindigkeit auf. Hierbei sind zwei Fälle zu unterscheiden:

1. Bereits eine einzelne Kante führt, wenn sie mit einer relativ zur Abtastweite geringen Geschwindigkeit bewegt wird, zu einer mehrdeutigen Detektorantwort (Abb. 3.64 A und B).
2. Bei zu großen Kantengeschwindigkeiten kommt ein strukturbedingtes Korrespondenzproblem hinzu. Dieses in realen Szenen häufig auftretende Problem ist in Abbildung 3.64 C und D anhand eines einfachen Objektes, das aus drei Linien besteht, dargestellt. Wird dieses Gitter um 4 Pixel nach rechts bewegt, dann detektieren die Richtungsdetektoren der mittleren Auflösung die richtige Bewegungsrichtung, doch die Richtungsdetektoren der feinen Auflösung detektieren aufgrund des Korrespondenzproblems eine lokale Scheinbewegung nach links.

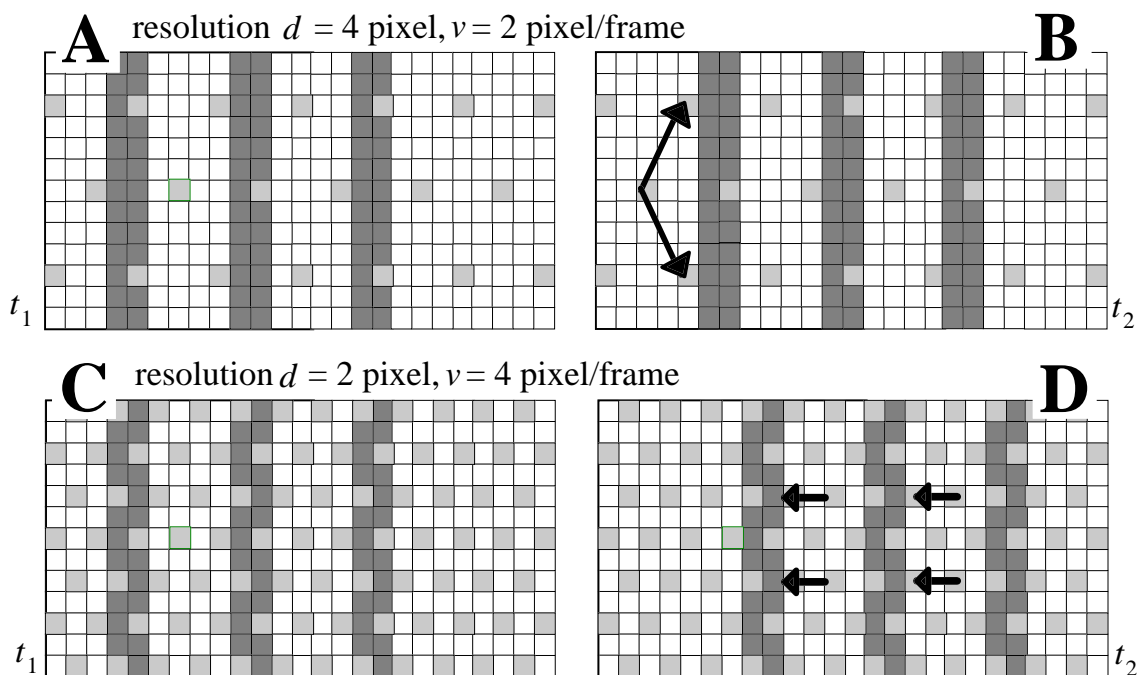


Abb. 3.64: Mehrdeutiges Ergebnis der Richtungsbestimmung mit Detektoren unterschiedlicher Auflösung. Dargestellt ist ein aus 3 Linien bestehendes Objekt, das sich in Richtung 0° mit einer Geschwindigkeit von $v = 2$ Pixel/Bild (A, B) bzw. $v = 4$ Pixel/Bild (C, D) bewegt. Die Richtungsdetektoren der mittleren Auflösung (A, B) zeigen eine falsche Bewegungsrichtung von 60° und 300° an. Die Richtungsdetektoren der feinen Auflösung (C, D) detektieren eine lokale Scheinbewegung in die entgegengesetzte Bewegungsrichtung.

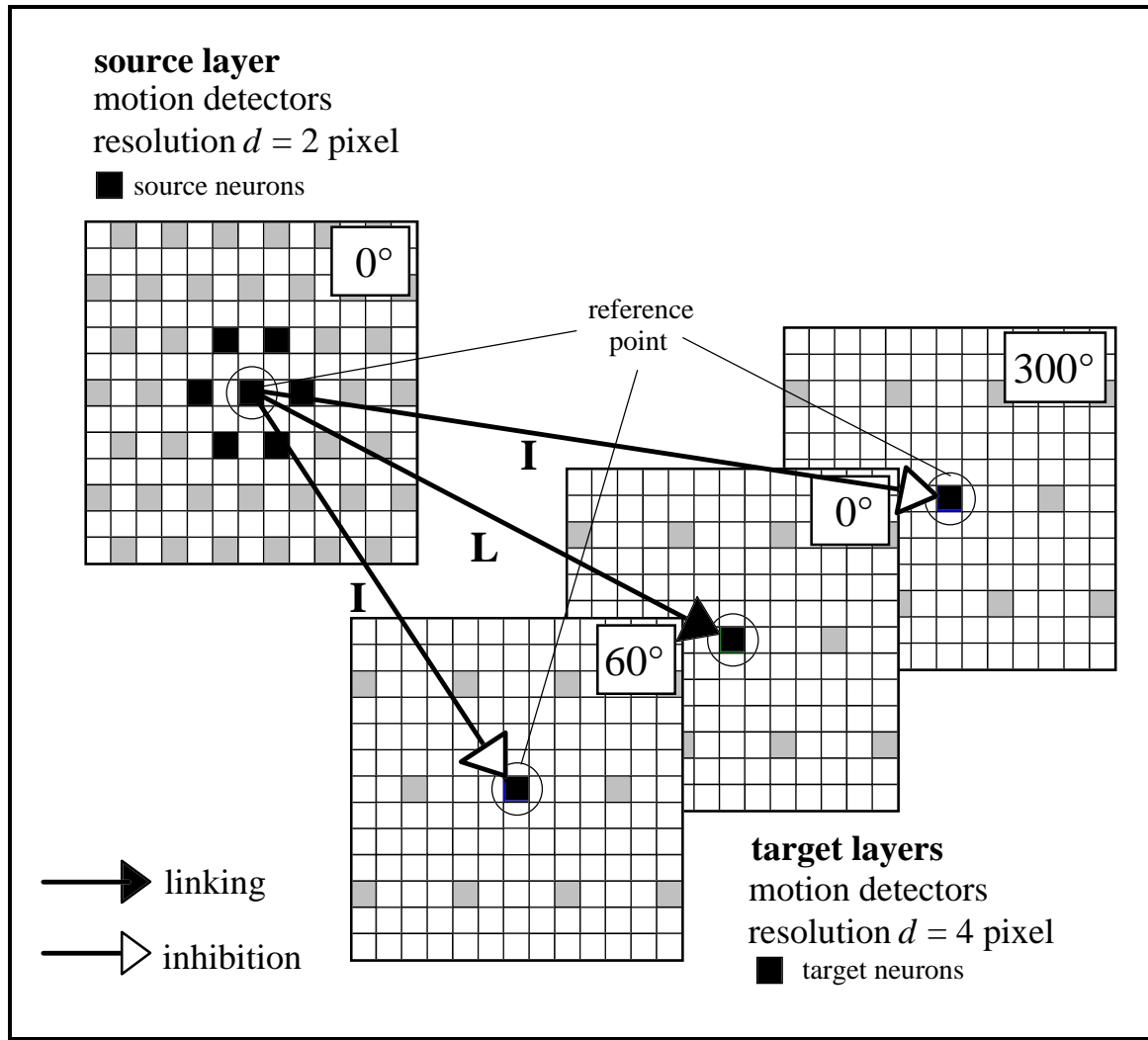


Abb. 3.65: Kopplung der Richtungsdetektoren unterschiedlicher Auflösung I. Die Abbildung zeigt anhand eines Beispiels die Kopplung zwischen Richtungsdetektoren der feinen und mittleren Auflösung. Aktivierte 0°-Richtungsdetektoren der feinen Auflösung unterstützen 0°-Richtungsdetektoren der mittleren Auflösung über unverzögerte Linking-Verbindungen. Sowohl die 60°-Richtungsdetektoren als auch die 300°-Detektoren werden inhibiert. Die Kopplung der anderen Richtungsdetektoren erfolgt in Analogie zu dem dargestellten Beispiel.

Aus diesen Gründen werden Richtungsdetektoren unterschiedlicher Auflösung lokal gekoppelt. Ziel ist es, die in einer bestimmten Auflösung falsch detektierte Bewegungsrichtung durch die richtige zu ersetzen.

Das erste Problem kann durch Einwirkung der Richtungsdetektoren der feinen auf die mittlere Auflösung behoben werden (Abb. 3.65). Die Richtungsdetektoren (i, j) der feinen Auflösung (Mf) unterstützen über unverzögerte Linking-Verbindungen $^{Mf}_w^{Mm}$ die Richtungsdetektoren (k, l) der mittleren Auflösung (Mm) mit gleicher Richtungspräferenz ϑ , wohingegen Richtungsdetektoren mit um 60° differierender Richtungspräferenz ϑ' von ihnen gehemmt werden (Gleichungen 3.50 und 3.51).

$${}^{Mf, \vartheta'=\vartheta} w_{ijkl}^{Mm, \vartheta} = \begin{cases} {}^{Mfm} V^L, & \text{falls } ((i == k) \wedge (j == l)) \vee \text{hexagon}_2 1 \\ 0, & \text{sonst} \end{cases} \quad (\text{Gl. 3.50})$$

$${}^{Mf, \vartheta'=\vartheta \pm 60} w_{ijkl}^{Mm, \vartheta} = \begin{cases} {}^{Mfm} V^I, & \text{falls } ((i == k) \wedge (j == l)) \vee \text{hexagon}_2 1 \\ 0, & \text{sonst} \end{cases} \quad (\text{Gl. 3.51})$$

Das zweite Problem kann durch Einwirkung der mittleren Auflösung auf die feine Auflösung behoben werden (Abb. 3.66, Gl. 3.52/3.53). Dieser Mechanismus wird durch die in Kapitel 2.1.2.3 beschriebene Untersuchung Ramachandrans unterstützt. Grobe Auflösungen dominieren die Bewegungswahrnehmung bei ausreichend großer Objektgeschwindigkeit und führen sogar dazu, daß diese Bewegungsrichtung auf die feine Struktur übertragen wird.

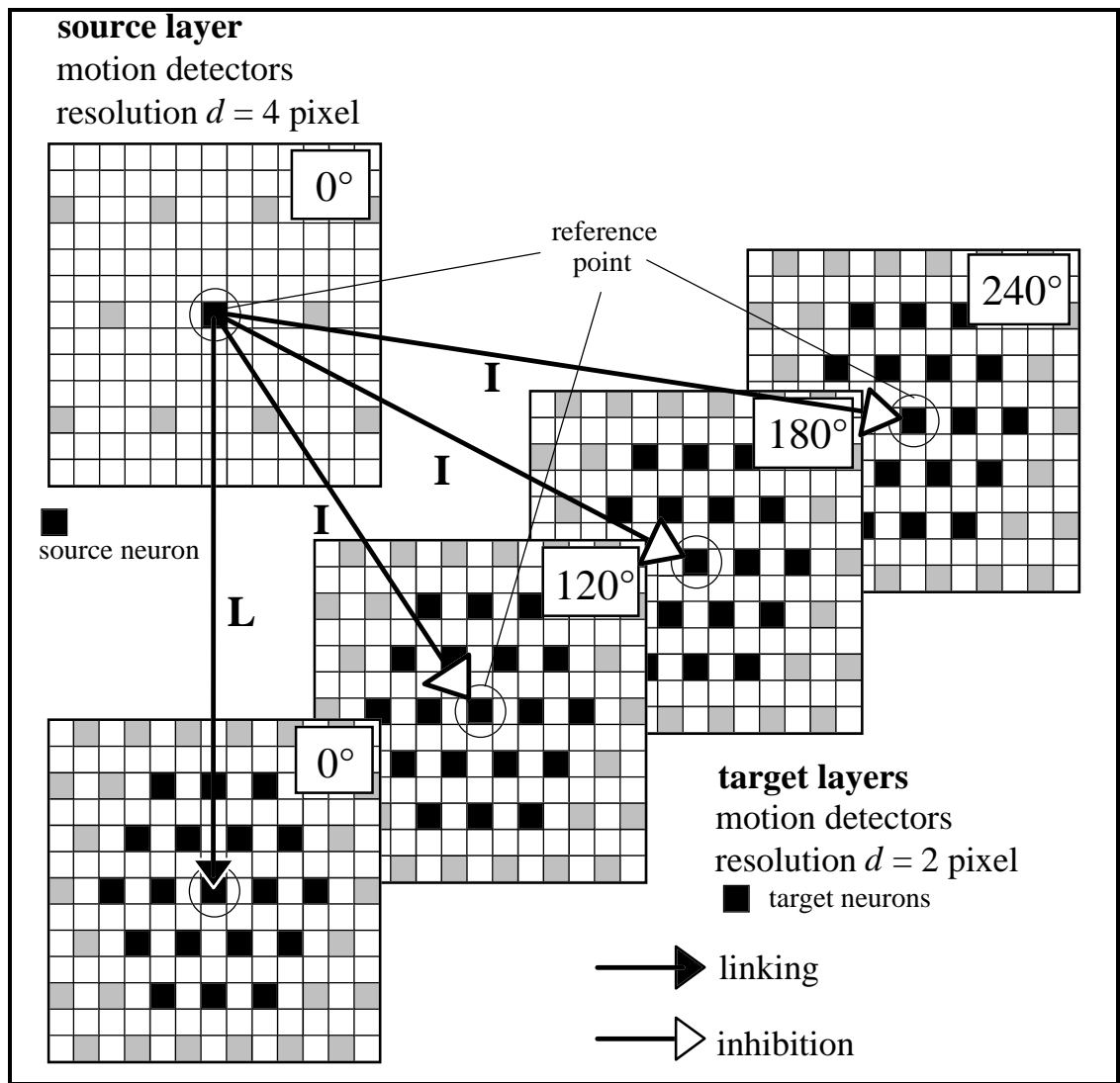


Abb. 3.66: Kopplung der Richtungsdetektoren unterschiedlicher Auflösung II. Die Abbildung zeigt anhand eines Beispiels die Kopplung zwischen Richtungsdetektoren der mittleren und feinen Auflösung. Aktivierte 0° -Richtungsdetektoren der mittleren Auflösung unterstützen 0° -Richtungsdetektoren der feinen Auflösung über unverzögerte Linking-Verbindungen. Die 120° -, 180° - und 240° -Richtungsdetektoren werden lokal inhibiert. Die Kopplung der anderen Richtungsdetektoren erfolgt in Analogie zu dem dargestellten Beispiel.

(Gl. 3.52)

$$w_{ijkl}^{Mm, \vartheta'=\vartheta, Mf, \vartheta} = \begin{cases} V^L, & \text{falls } ((i=k) \wedge (j=l)) \vee \text{hexagon}_2 1 \vee \text{hexagon}_2 2 \\ 0, & \text{sonst} \end{cases}$$

(Gl. 3.53)

$$w_{ijkl}^{Mm, \vartheta+60^\circ < \vartheta' < \vartheta-60^\circ, Mf, \vartheta} = \begin{cases} V^I, & \text{falls } ((i=k) \wedge (j=l)) \vee \text{hexagon}_2 1 \vee \text{hexagon}_2 2 \\ 0, & \text{sonst} \end{cases}$$

Die Kopplung findet per se nur in überlappenden Abtastbereichen statt. Die Interaktion erfolgt über Linking-Verbindungen, da Feeding-Verbindungen durch eine Aktivitätsverschmierung die Struktur der rezeptiven Felder der Richtungsdetektoren stören würden. Feeding-Verbindungen würden bei ausreichend starker Aktivierung eine Bewegung in statischen Bildbereichen vortäuschen.

Die Inhibition sorgt für eine Unterdrückung der falsch antwortenden Richtungsdetektoren. Sie wird so gewählt, daß die Einwirkung der mittleren auf die feine Auflösung eine möglicherweise auf dem ersten Problem beruhende richtige Detektorantwort in der feinen Auflösung nicht zunichte macht.

Die lokale Reichweite der Interaktion wurde willkürlich festgelegt. Je größer diese Reichweite ist, umso besser können die Richtungsdetektoren der anderen Auflösung beeinflußt werden. Da die Kopplung isotrop erfolgt, besteht bei zu großen Reichweiten allerdings die Gefahr, die Antworten der Richtungsdetektoren anderer Objekte fehlerhaft zu verändern.

Für die Linking-Verbindungen wurden kleine Gewichtswerte festgelegt, damit eine Verschiebung in der Aktivitätsverteilung der Empfängerneuronen nur bei starker Aktivierung der Sendeneuronen erfolgt. Die zusätzlich verwendeten Parameter sind in Tabelle 3.7 zusammengefaßt. Die anderen Parameter bleiben unverändert.

Linking-Gewicht (mittlere Auflösung) $M_{fm} V^L$	0,5
Linking-Gewicht (feine Auflösung) $M_{mf} V^L$	0,5
Inhibitions-Gewicht (mittlere Auflösung) $M_{fm} V^I$	7,5
Inhibitions-Gewicht (feine Auflösung) $M_{mf} V^I$	7,5
Zeitkonstante des Linking-Eingangs (alle Auflösungen) τ^L	5 Bin
Zeitkonstante des Linking-Eingangs (alle Auflösungen) τ^I	7,5 Bin

Tab. 3.7: Parameter der Kopplung der Richtungsdetektoren unterschiedlicher Auflösung.

3.3.2 Simulationen zur Auflösungskopplung

3.3.2.1 Simulationsergebnisse mit künstlichen Szenen

Die Auswirkung der Interaktion zwischen Richtungsdetektoren unterschiedlicher Auflösung wird zunächst anhand von drei künstlichen Szenen und dann mittels der realen Szene „Durlacher Tor“ demonstriert. Alle Simulationen beinhalten *motion capture* zwischen Richtungsdetektoren unterschiedlicher Auflösung.

Gitterreiz

Der Stimulus der ersten Simulation besteht aus dem in Abbildung 3.64 dargestellten Gitter. Die Abbildung 3.67 zeigt die über die Simulationsdauer von $t = 1509$ Bin kumulierte Aktivität der ONOFF-Richtungsdetektoren der beiden Auflösungen für die Gittergeschwindigkeiten $v = 2$ Pixel/Bild und $v = 4$ Pixel/Bild. Es wurden 20×20 Abtastpunkte in der feinen und 10×10 Abtastpunkte in der mittleren Auflösung verwendet. Bewegt sich das Gitter mit der für die jeweilige Auflösung optimalen Geschwindigkeit, wird sowohl mit als auch ohne Interaktion die richtige Bewegungsrichtung detektiert (Abb. 3.67 A, D).

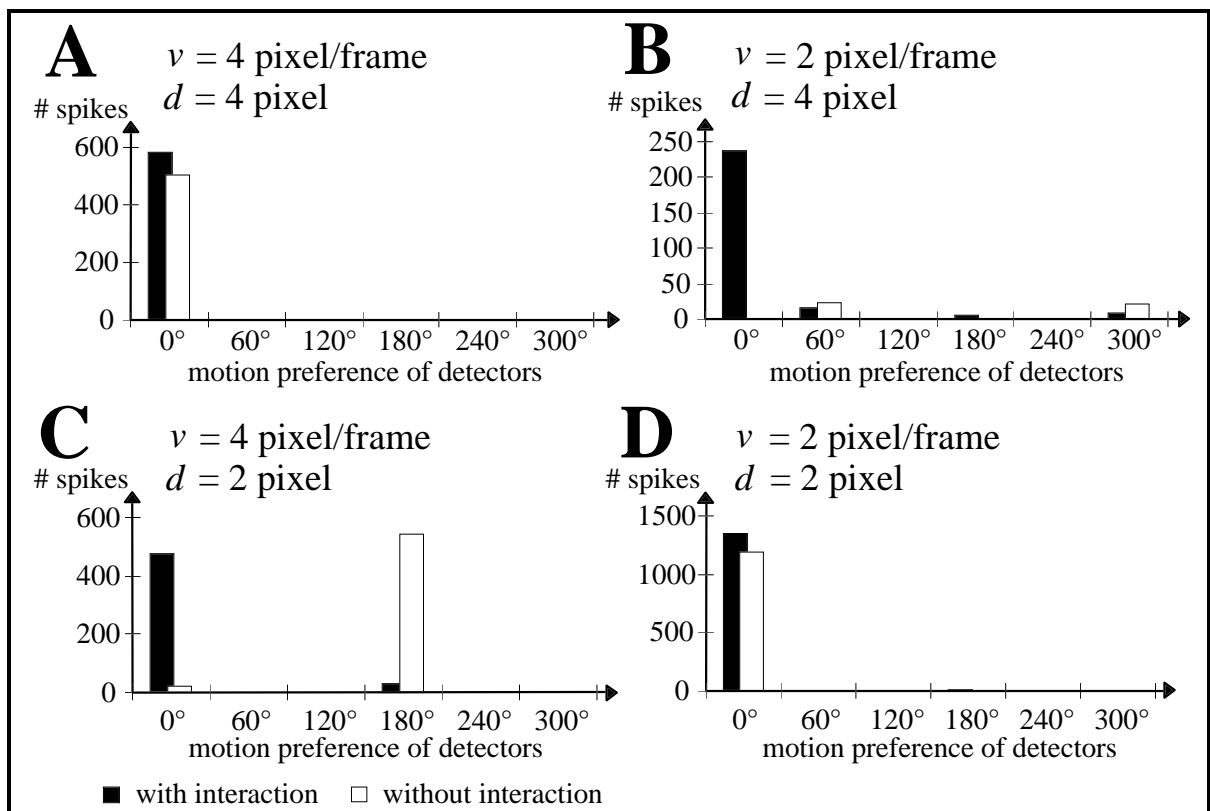


Abb. 3.67: Simulationsergebnis „Gitterreiz“. Dargestellt ist die gemittelte Aktivität der ONOFF- und OFFON-Richtungsdetektoren der Auflösungen $d = 2$ Pixel (C, D) und $d = 4$ Pixel (A, B) mit und ohne Interaktion für die Geschwindigkeiten $v = 2$ Pixel/Bild (B, D) und $v = 4$ Pixel/Bild (A, C) (Gitterbewegungsrichtung 0°). A und D: Das Gitter bewegt sich mit der für die jeweilige Auflösung optimalen Geschwindigkeit. Die Bewegungsrichtung wird richtig detektiert. B: Das Gitter bewegt sich zu langsam für die Richtungsdetektoren der mittleren Auflösung, wodurch es ohne Interaktion zu einem mehrdeutigen Ergebnis der Bewegungsrichtungsdetektion kommt. Mit Interaktion wird die korrekte Bewegungsrichtung detektiert. C: Der Gitterreiz bewegt sich zu schnell, wodurch es zu einem stimulusbedingten Korrespondenzproblem kommt. Mit Interaktion werden die Richtungsdetektoren der feinen Auflösung von den Richtungsdetektoren der mittleren Auflösung dominiert.

Die Richtungsdetektoren der mittleren Auflösung antworten bei einer Gittergeschwindigkeit von $v = 2$ Pixel/Bild fehlerhaft (Abb. 3.67 B). Ohne Interaktion sind keine 0° -Richtungsdetektoren aktiv. Wie bereits oben aufgeführt, führt diese Geschwindigkeit zu einer schwachen und mehrdeutigen Detektorantwort (60° und 300°). Werden Interaktionen zugelassen, dann dominieren die Detektoren, die die richtige Gitterbewegungsrichtung repräsentieren.

Die Detektoren der feinen Auflösung detektieren bei einer Geschwindigkeit von $v = 4$ Pixel/Bild ohne Interaktion eine Scheinbewegung in entgegengesetzter Richtung (Abb. 3.67 C). Mit Interaktion dominieren hier die 0° -Richtungsdetektoren. In diesem Beispiel wird *motion capture* unter Verwendung unterschiedlicher Auflösungen gezeigt.

Richtungscharakteristiken

In der zweiten Simulation werden die Richtungscharakteristiken der Modellneuronen der mittleren und der feinen Auflösung untersucht. Dazu wird in Analogie zu Kapitel 3.2.2.1 eine einfache Intensitätskontrastkante als Stimulus benutzt. In Abbildung 3.46 zeigte sich eine mehrdeutige Antwort der Richtungsdetektoren der mittleren Auflösung bei Kantengeschwindigkeiten von $v = 2$ Pixel/Bild und $v = 3$ Pixel/Bild. Dieser Geschwindigkeitsbereich wurde hingegen eindeutig durch die Richtungsdetektoren der feinen Auflösung abgedeckt, wie in Abbildung 3.44 gezeigt werden konnte.

Die Interaktion zwischen den Richtungsdetektoren der beiden Auflösungen führt auch in der mittleren Auflösung zu einem eindeutigen Ergebnis der Bewegungsdetektion. Die Abbildung 3.68 zeigt die Richtungscharakteristiken der Richtungsdetektoren der mittleren Auflösung für sechs verschiedene Kantengeschwindigkeiten. Darin wird deutlich, daß für $v = 2$ Pixel/Bild und $v = 3$ Pixel/Bild die Richtungsdetektion eindeutig ist. Die stärkere Aktivierung der Richtungsdetektoren der feinen Auflösung setzt sich gegen die schwächere der Richtungsdetektoren der mittleren Auflösung durch. Für Geschwindigkeiten von $v = 4$ Pixel/Bild und $v = 5$ Pixel/Bild werden die Aktivitäten der Richtungsdetektoren der mittleren Auflösung etwa um 25% höher, ohne dabei die Halbwertsbreite der Aktivitätsverteilungen zu verändern. Die Interaktion zeigt keinen signifikanten Einfluß bei Kantengeschwindigkeiten von $v = 6$ Pixel/Bild. Die Unsymmetrie in den Richtungscharakteristiken bei kleinen Geschwindigkeiten resultiert aus Rand- und Quantisierungseffekten, die durch die neuronalen Kopplungen zusätzlich verstärkt werden. Die Unterschiede zwischen den Amplituden der 60° -, 120° -, 240° - und 300° Richtungsdetektoren zu denen der 0° - und 180° -Richtungsdetektoren resultieren aus den unterschiedlichen Abständen zwischen benachbarten Abtastpunkten (z.B. in Richtung 0° 2 Pixel und in Richtung 60° $\sqrt{5}$ Pixel).

Die Abbildung 3.69 zeigt ein qualitatives Gleichbleiben der Richtungscharakteristiken der Richtungsdetektoren der feinen Auflösung für Geschwindigkeiten von $v = 2$ Pixel/Bild und $v = 3$ Pixel/Bild. Die Aktivität bei Stimulation mit einer Kante, die sich mit $v = 2$ Pixel/Bild bewegt, ist ungefähr um 20% höher als die ohne Interaktion zwischen den Richtungsdetektoren beider Auflösungen. Die etwas schwache Aktivierung bei Kantengeschwindigkeiten von $v = 3$ Pixel/Bild ohne Interaktion wird mit Interaktion um 60% erhöht. Für Geschwindigkeiten von $v = 4$ Pixel/Bild, $v = 5$ Pixel/Bild und $v = 6$ Pixel/Bild zeigen die Richtungsdetektoren der feinen Auflösung mit Interaktion eine schwache, aber eindeutige Aktivierung.

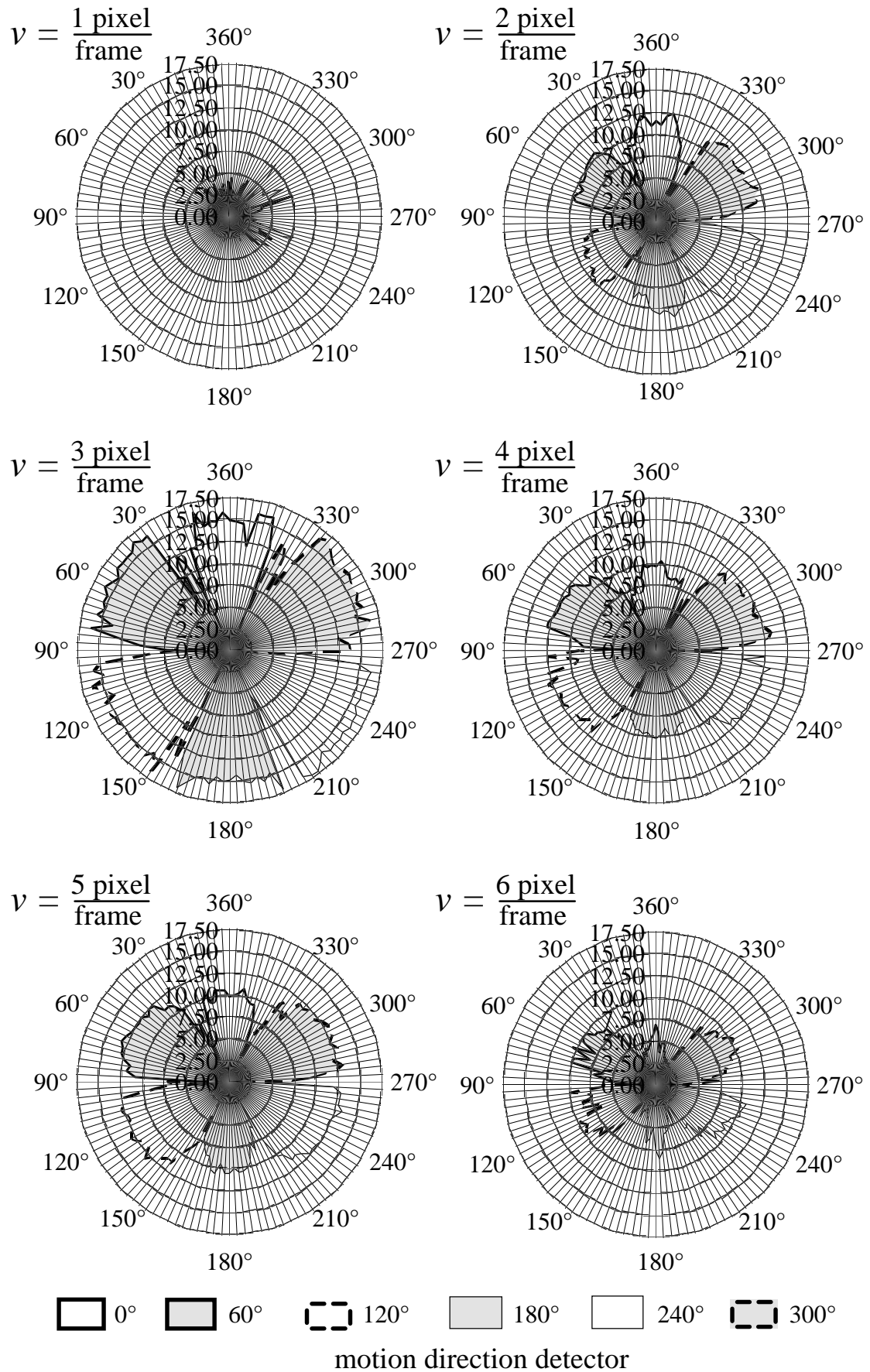


Abb. 3.68: Richtungscharakteristiken der Richtungsdetektoren der mittleren Auflösung mit Interaktion zwischen den Auflösungsschichten.

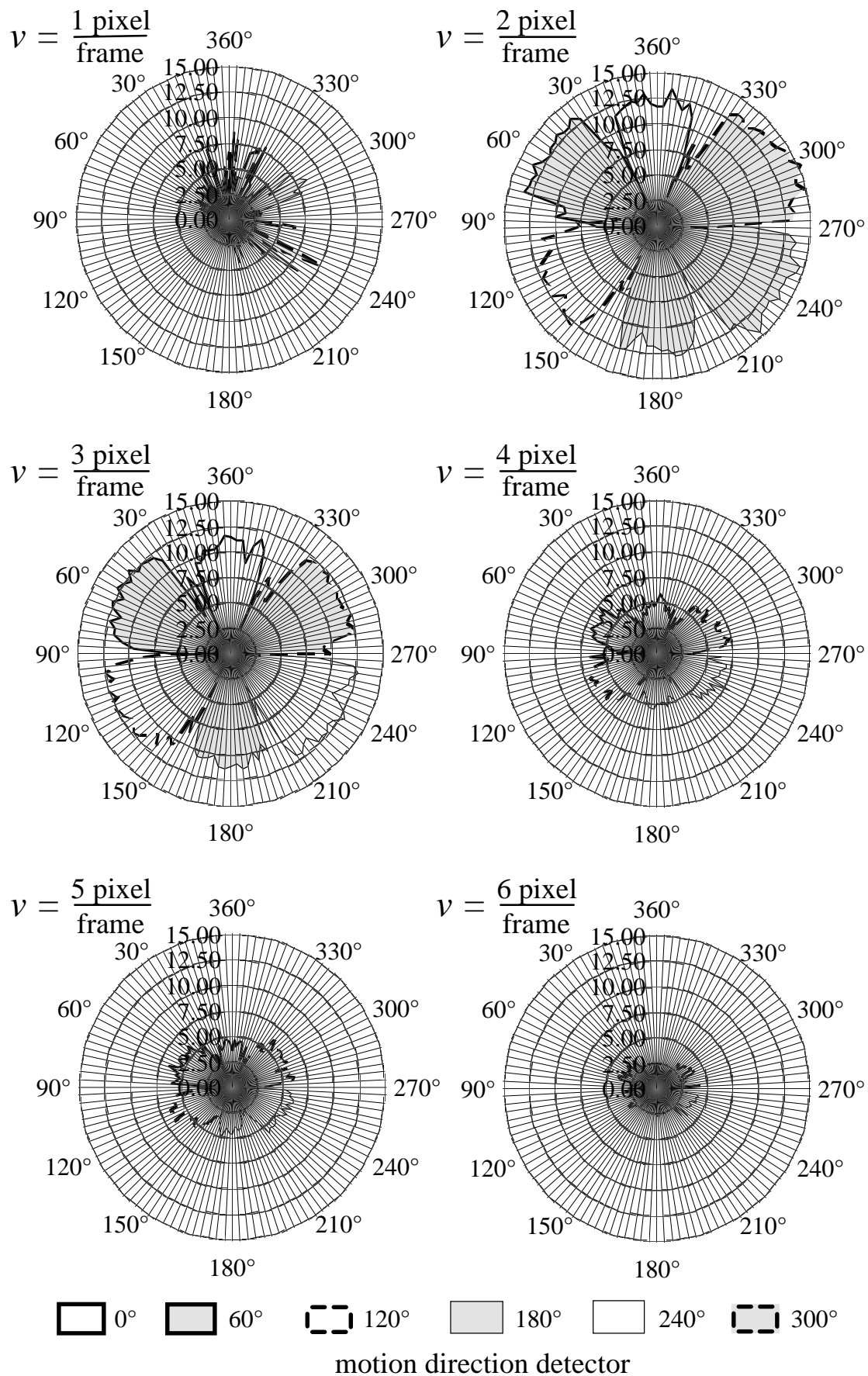


Abb. 3.69: Richtungscharakteristiken der Richtungsdetektoren der feinen Auflösung mit Interaktion zwischen den Auflösungsschichten.

Motion Capture bei „Fleckenmuster“

Die dritte Simulation mit künstlicher Eingabe zeigt den von Ramachandran und Cavanagh 1987 beschriebenen Effekt. Dazu wurden dem gleichen neuronalen Netz 128 Pixel x 128 Pixel große Bilder mit einem sinusförmigen Gitter, dessen Grauwertverlauf $gw(x)$ sich nach Gleichung 3.54 ergibt, und 100 darin zufällig verteilten 3 Pixel x 3 Pixel großen Flecken mit Grauwert 100 als Eingabe zugeführt. Die Flecken bewegten sich mit $v = 2$ Pixel/Bild in Richtung 0° . Bei Verlassen des Bildbereiches wurden die Flecken auf der entgegengesetzten Seite wieder eingesetzt. Das Gitter bewegte sich gleichmäßig mit $v = 0$ Pixel/Bild bis $v = 6$ Pixel/Bild in Richtung 180° . Jede Simulation beinhaltete 1568 Zeitschritte.

$$gw(x) = 100 \cdot \sin(x/k) + 100 \quad (\text{Gl. 3.54})$$

Das Simulationsergebnis ist in Abbildung 3.70 dargestellt. Es wurden drei verschiedene Werte für k verwendet: $k_1 = 4$ Pixel (A, B), $k_2 = 5$ Pixel (C, D) und $k_3 = 6$ Pixel (E, F), so daß sich drei Wellenlängen λ mit $\lambda_1 = 26$ Pixel, $\lambda_2 = 32$ Pixel bzw. $\lambda_3 = 38$ Pixel ergaben. In den Diagrammen der Abbildung 3.65 sind die mittleren kumulierten Aktivitäten der 0° - und 180° -ON- und OFF-Richtungsdetektoren der feinen Auflösung mit ihren Standardabweichungen aufgetragen. Die Mittelwerte sind durch Subtraktion um die Aktivität bereinigt, die das Gitter alleine verursacht. Mit zunehmender Gittergeschwindigkeit fällt die Aktivität der 0° -Richtungsdetektoren, die Aktivität der 180° -Richtungsdetektoren steigt. Diese Abhängigkeit von der Gittergeschwindigkeit liegt natürlich in der Aktivierung der Richtungsdetektoren der mittleren Auflösung begründet. Bei kleinen Geschwindigkeiten $v \leq 2$ Pixel/Bild ist diese Aktivität noch zu gering. Wird die Wellenlänge noch größer gewählt, wird der Effekt geringer, da die divergente Verschaltung der mittleren Auflösung auf die feine Auflösung nicht mehr ausreicht und der Grauwertgradient des Gitters zu klein ist.

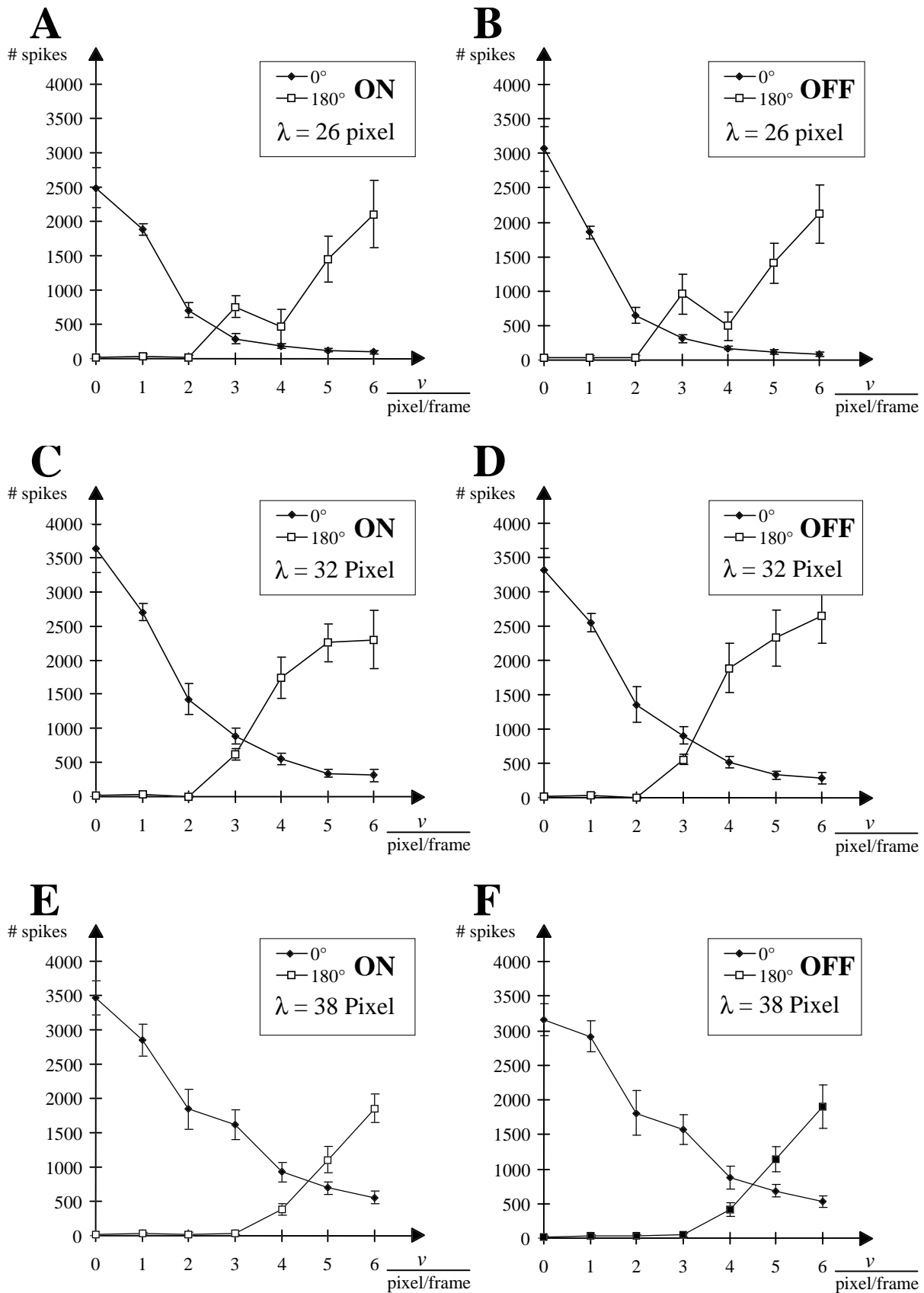


Abb. 3.70: Simulationsergebnis „motion capture“ zwischen Richtungsdetektoren unterschiedlicher Auflösung. Für drei verschiedene Gitterwellenlängen λ (A, B; C, D; E, F) sind jeweils die gemittelten Aktivitäten der 0° - und 180° - ON- (A, C, E) und OFF- (B, D, F) Richtungsdetektoren der feinen Auflösung und deren doppelte Standardabweichung in Abhängigkeit von der Gittergeschwindigkeit v aufgetragen. Die Mittelwerte wurden um die Aktivitäten bereinigt, die das Gitter alleine verursacht. Das Simulationsergebnis zeigt den von Ramachandran beobachteten psychophysischen Effekt (siehe Text).

3.3.2.2 Simulationsergebnisse mit realen Szenen

Die Auswirkungen der Interaktion zwischen den Detektoren bei der Bewegungsrichtungsdetektion in der realen Szene „Durlacher Tor“ zeigen die nachfolgenden Abbildungen. Es wurden auch hier die feine ($d = 2$ Pixel) und die mittlere ($d = 4$ Pixel) Auflösung verwendet. Durch die Verwendung von zwei Auflösungen reduziert sich der zu verarbeitende Ausschnitt auf 32×32 Abtastpunkte. Die Simulationszeit betrug $t = 1536$ Bin.

Die Abbildung 3.71 zeigt das Simulationsergebnis von fünf aufeinanderfolgenden Bildern der Szene „Durlacher Tor“. Die Fahrzeuge bewegen sich mit einer Geschwindigkeit von $v = 1$ Pixel/Bild bis $v = 2$ Pixel/Bild auf einer Kurve von links unten nach rechts oben durch den Bildausschnitt hindurch. Der eingezeichnete Rahmen stellt den Verarbeitungsbereich der feinen Auflösung dar. In der linken Spalte ist die Aktivität der Richtungsdetektoren der feinen Auflösung dargestellt. Es kommt zu einer starken Aktivität der 60° - und zu einer geringen Aktivität der 0° -ONOFF-Richtungsdetektoren. Die Abbildungen 3.72 und 3.73 zeigen die kumulierten Aktivitäten der Richtungsdetektoren bzw. der ONOFF-Richtungsdetektoren über die gesamte Simulationsdauer auf. Insbesondere bei den ONOFF-Richtungsdetektoren ist eine dominante Antwort für die 60° -Richtung zu sehen. Dies entspricht der Hauptbewegungsrichtung der Fahrzeuge.

Im Gegensatz dazu steht die Aktivität der Richtungsdetektoren der mittleren Auflösungen. Sie ist exemplarisch in der mittleren Spalte der Abbildung 3.71 dargestellt. Aufgrund der niedrigen Fahrzeuggeschwindigkeit ist die Aktivität gering. Die 0° -ONOFF-Richtungsdetektoren zeigen die stärkste Aktivität, während die 60° -ONOFF-Richtungsdetektoren kaum aktiv sind. Die Richtungsdetektoren der mittleren Auflösung spiegeln nicht die richtige Objektbewegungsrichtung wider.

Mit der Interaktion zwischen den Richtungsdetektoren unterschiedlicher Auflösungen ändert sich das Verhalten der Richtungsdetektoren der mittleren Auflösung. Die 0° -Richtungsdetektoren der mittleren Auflösung werden durch die stark aktiven 60° -Richtungsdetektoren der feinen Auflösung gehemmt. Die 60° -Richtungsdetektoren der mittleren Auflösung werden hingegen unterstützt. Dadurch wird nun auch in der mittleren Auflösung, wenn auch nur durch eine geringe Aktivität, die richtige Objektbewegungsrichtung angezeigt.

Die Abbildungen 3.74 und 3.75 zeigen die unterschiedlichen Antwortverhalten der Richtungsdetektoren bzw. ONOFF-Richtungsdetektoren der mittleren Auflösung mit und ohne Interaktion auf. Bereits die Aktivitäten der Richtungsdetektoren zeigen eine Verschiebung des Aktivitätsmaximums zu den 60° -Richtungsdetektoren (Abb. 3.74). Durch die Kopplung von ON- und OFF-Pfad wird dieser Effekt noch verstärkt, so daß es von einer detektierten Bewegung in Richtung 0° zu einer in Richtung 60° kommt.

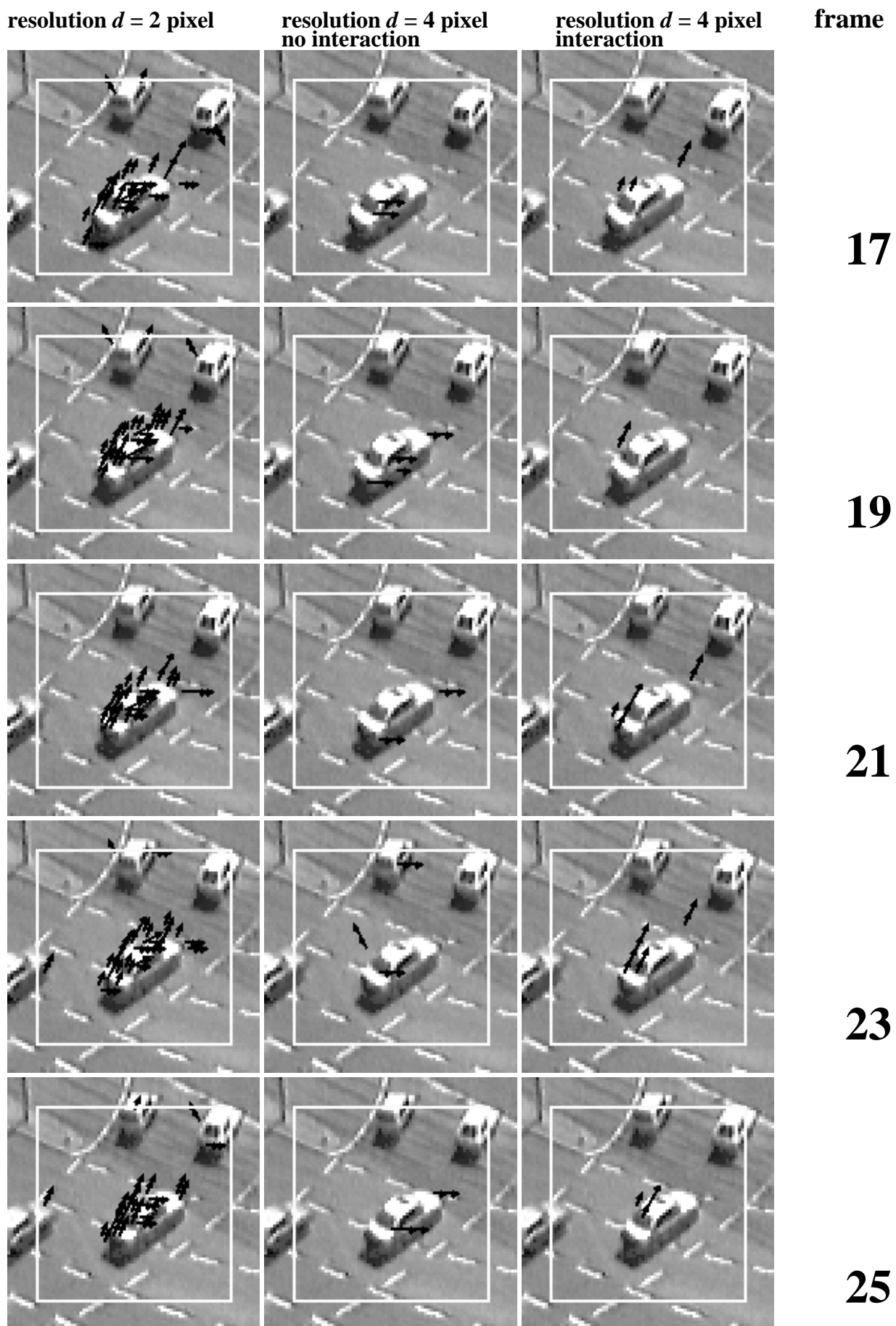


Abb. 3.71: Interaktion der Richtungsdetektoren unterschiedlicher Auflösung. Siehe Text.

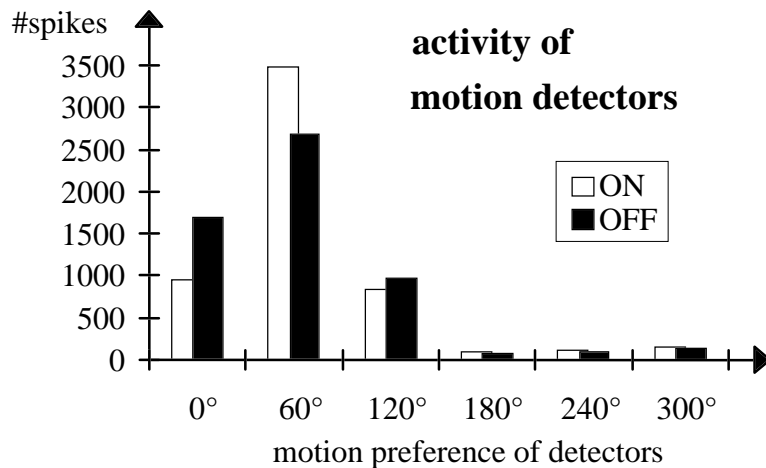


Abb. 3.72: Aktivität der Richtungsdetektoren der feinen Auflösung (Sequenz „Durlacher Tor“). Dargestellt ist die Anzahl der Aktionspotentiale der Richtungsdetektoren innerhalb der Simulationsdauer von 1536 Bin unter Verwendung der Sequenz „Durlacher Tor“ als Eingabe (32 x 32 Abtastpunkte, Abtastbereich ($x=80$ Pixel, $y=64$ Pixel) bis ($x=144$ Pixel, $y=128$ Pixel)).

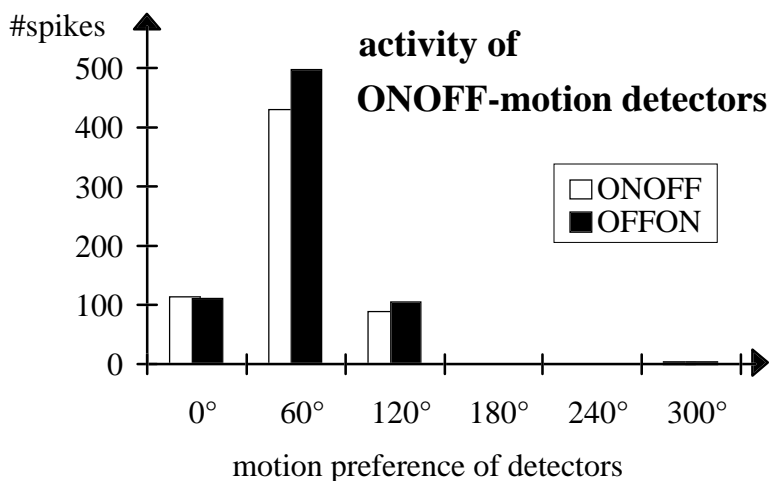


Abb. 3.73: Aktivität der ONOFF-Richtungsdetektoren der feinen Auflösung (Sequenz „Durlacher Tor“). Dargestellt ist die Anzahl der Aktionspotentiale der ONOFF-Richtungsdetektoren innerhalb der Simulationsdauer von 1536 Bin unter Verwendung der Sequenz „Durlacher Tor“ als Eingabe (32 x 32 Abtastpunkte, Abtastbereich (80 Pixel, 64 Pixel) bis (144 Pixel, 128 Pixel)). Im Gegensatz zu den vorgeschalteten Richtungsdetektoren weisen die ONOFF-Richtungsdetektoren eine deutlich geringere Gesamtaktivität auf.

Die beiden Abbildungen zeigen ebenso die Notwendigkeit der Interaktion zwischen den Richtungsdetektoren und nicht die zwischen den ONOFF-Richtungsdetektoren. Die ONOFF-Richtungsdetektoren zeigen eine deutlich geringere Aktivität. Insbesondere die 60°-ONOFF-Richtungsdetektoren (Abb. 3.75) sind ohne Interaktion nahezu inaktiv. In den ONOFF-Richtungsdetektorschichten wurde die notwendige Bewegungsinformation, die durch eine schwache Aktivität getragen wurde, bereits herausgefiltert. Die 60°-Richtungsdetektoren zeigen hingegen eine mittlere Aktivität. Dadurch ist der Einsatz von unverzögerten Linking-Kopplungen zwischen den Richtungsdetektoren sinnvoller als zwischen ONOFF-Richtungsdetektoren. Die Abbildung 3.76 zeigt die Aktivität der ONOFF-Richtungsdetektoren mit Kopplung.

Das Beispiel zeigt die Interaktion zwischen Richtungsdetektoren unterschiedlicher Auflösungen. Die Interaktion ist nur zwischen überlappenden Abtastbereichen möglich. Nachteilig an diesem Mechanismus ist die hohe Anzahl an Modellneuronen. Bei Verwendung eines einzigen Acceleratorsystems kann dieser Mechanismus mit Rücksicht auf einen ausreichend großen Abtastbereich nicht verwendet werden.

Abb. 3.74: Aktivität der Richtungsdetektoren der mittleren Auflösung (Sequenz „Durlacher Tor“). Dargestellt ist die Anzahl der Aktionspotentiale der Richtungsdetektoren innerhalb der Simulationsdauer von 1536 Bin unter Verwendung der Sequenz „Durlacher Tor“ als Eingabe mit und ohne Interaktion mit den Richtungsdetektoren der feinen Auflösung (16 x 16 Abtastpunkte, Verarbeitungsbereich (80 Pixel, 64 Pixel) bis (144 Pixel, 128 Pixel)). Das Aktivitätsmaximum verschiebt sich von den 0°-Richtungsdetektoren zu den 60° Richtungsdetektoren.

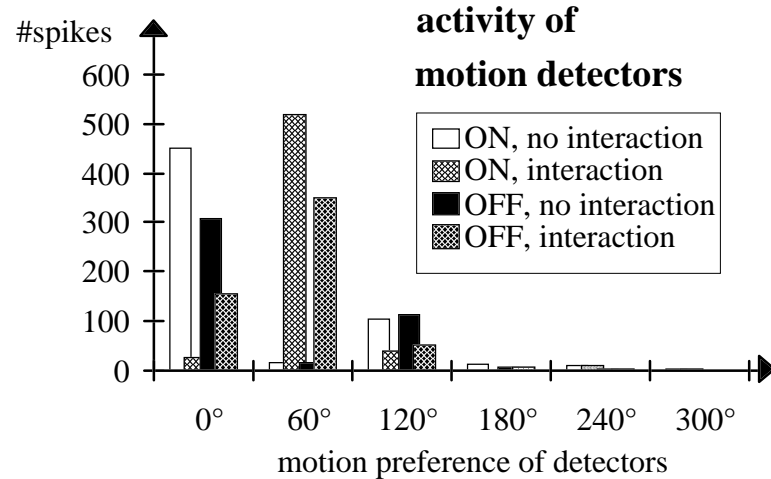


Abb. 3.75: Aktivität der ONOFF-Richtungsdetektoren der mittleren Auflösung (Sequenz „Durlacher Tor“). Dargestellt ist die Anzahl der Aktionspotentiale der ONOFF-Richtungsdetektoren innerhalb der Simulationsdauer von 1536 Bin unter Verwendung der Sequenz „Durlacher Tor“ als Eingabe mit und ohne Interaktion mit den Richtungsdetektoren der feinen Auflösung (16 x 16 Abtastpunkte, Verarbeitungsbereich (80 Pixel, 64 Pixel) bis (144 Pixel, 128 Pixel)). Ohne Interaktion sind nur die 0°-ONOFF-Richtungsdetektoren aktiv, mit Interaktion nur die 60°-Detektoren.

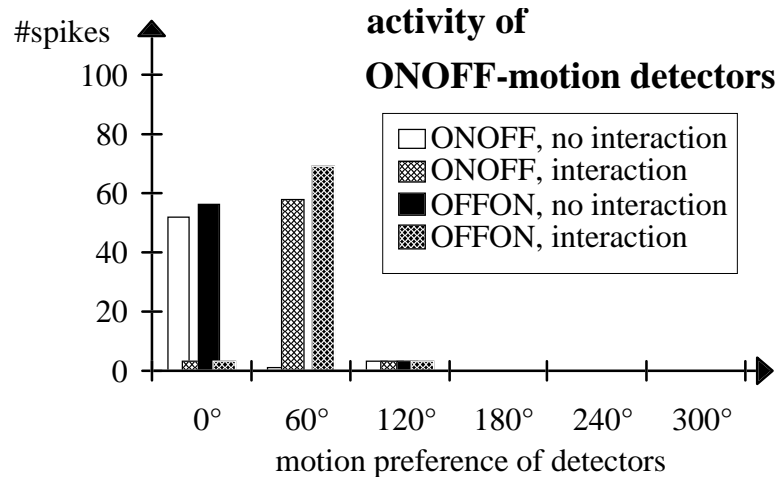
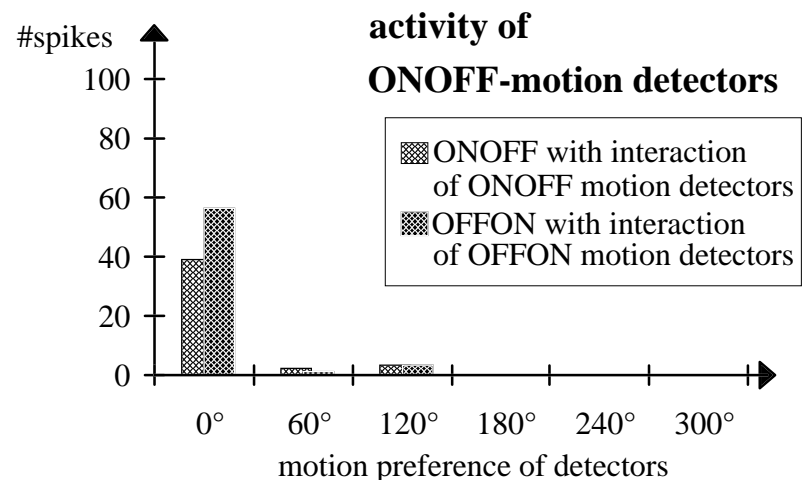


Abb. 3.76: Aktivität der ONOFF-Richtungsdetektoren der mittleren Auflösung mit Kopplung der ONOFF-Richtungsdetektoren (Sequenz „Durlacher Tor“). Dargestellt ist die Anzahl der Aktionspotentiale der ONOFF-Richtungsdetektoren innerhalb der Simulationsdauer von 1536 Bin unter Verwendung der Sequenz „Durlacher Tor“ als Eingabe mit Interaktion zwischen den ONOFF-Richtungsdetektoren (16 x 16 Abtastpunkte, Verarbeitungsbereich (80 Pixel, 64 Pixel) bis (144 Pixel, 128 Pixel)). Trotz Interaktion sind nur die 0°-ONOFF-Richtungsdetektoren aktiv.



3.4 Einbindung der Linienenden in die Bewegungsverarbeitung

Wie in der Einleitung bereits dargestellt wurde, sind Kontraste in den Merkmalsbereichen „Intensität“, „Bewegung“ und „Farbe“ Indizien für Objektgrenzen. Die lokalen Kontraste sind Bildbereiche hohen Informationsgehaltes.

Bildbereiche mit genügend großem Intensitätskontrast werden in diesem Modell durch aktive Ganglienzellen markiert. Aus der so gewonnenen Objektkontur können wiederum Bereiche hohen Informationsgehaltes, das sind Bereiche mit großen Orientierungsänderungen (Ecken) und Konturverzweigungen oder -unterbrechungen (T-Stücke bzw. Linienenden), bestimmt werden. Diese sogenannten Vertizes (Biederman 1987, Hummel und Biederman 1992, Heitger et al. 1992) sind nicht nur für die Objektklassifizierung, sondern auch für die Bewegungsanalyse von Bedeutung (Hildreth 1984, Nakayama und Silverman 1988a, b). Je weniger Merkmale das Bewegungssystem verarbeiten muß, um so geringer ist das Korrespondenzproblem und um so sicherer ist die Bestimmung der Bewegungsrichtung. Deswegen kann die Bewegungsrichtung eines Objektes besonders gut an seinen Vertizes ermittelt werden. Die im Grundlagenteil beschriebene *barberpole illusion* zeigt einerseits die biologische Relevanz der Einbindung der Linienenden in die Bewegungsverarbeitung, andererseits - mit Hinblick auf die Detektion der physikalischen Bewegungsrichtung - auch die Grenzen in der technischen Bildverarbeitung.

In diesem Kapitel wird ein impulskodierendes neuronales Netzwerk (Schott et al. 1997, Schott und Eckhorn 1998) vorgestellt, das eine mögliche Erklärung der *barberpole illusion* liefert. Die von dem neuronalen Netzwerk extrahierten Linienenden werden im Sinne von Nakayama und Silverman (1988b) als intrinsisch angenommen. Der Einfluß der Tiefenwahrnehmung wird aufgrund der immensen Anzahl an Modellneuronen, die zur Simulation erforderlich wären, nicht untersucht.

3.4.1 Gesamtmodell zur Einbindung der Linienenden in die Bewegungsverarbeitung

Die nachfolgende Abbildung 3.77 zeigt als Überblick das Gesamtmodell des Transientensystems. Der Hauptverarbeitungspfad (Abb. 3.77 1, 2, 5b, 6 und 7) extrahiert die Bewegungsrichtung der Kontur (5b) unter Verwendung zeitlicher Transienten (2) der lokalen Intensitätskontraste (1), um aus dieser den Bewegungskontrast (6) zur Objektdefinition oder zur Aufmerksamkeitssteuerung zu ermitteln und eine Tiefenschätzung (7) durch Bewegungsparallaxe (Saam et al. 1997) machen zu können. Mit dem Hauptverarbeitungspfad allein läßt sich die *barberpole illusion* nicht erklären. Hierzu ist zusätzlich der Nebenpfad (Abb. 3.77 3, 4 und 5a) erforderlich. Aus den zeitlichen Transienten lokaler Intensitätskontraste (2) werden zunächst kurze, zusammenhängende Kontursegmente und deren jeweilige lokale Orientierung bestimmt (3). Dieser Schritt ist nötig, um eine sichere Detektion der scharfen Krümmungen und der Linienenden zu gewährleisten (4). An die Detektion der Linienenden schließt sich die Bestimmung ihrer jeweiligen Bewegungsrichtung an (5a). Die Bewegungsrichtung der Linienenden wird mit einer neuronalen Verschaltung implementiert, die der Verschaltung zur Richtungsdetektion der Kontur sehr ähnlich ist. Das Ergebnis der Linienendenbewegungsrichtungsdetektion (LEBRD) wird dann zur Beeinflussung der Konturbewegungsrichtungsdetektion (KBRD) verwendet. Ferner kann es noch über eine Rückkopplung zur Verbesserung der Detektionseigenschaften der Linienendendetektoren beitragen. Die Stufen 3, 4 und 5a des Transientensystems werden nachfolgend genauer erklärt.

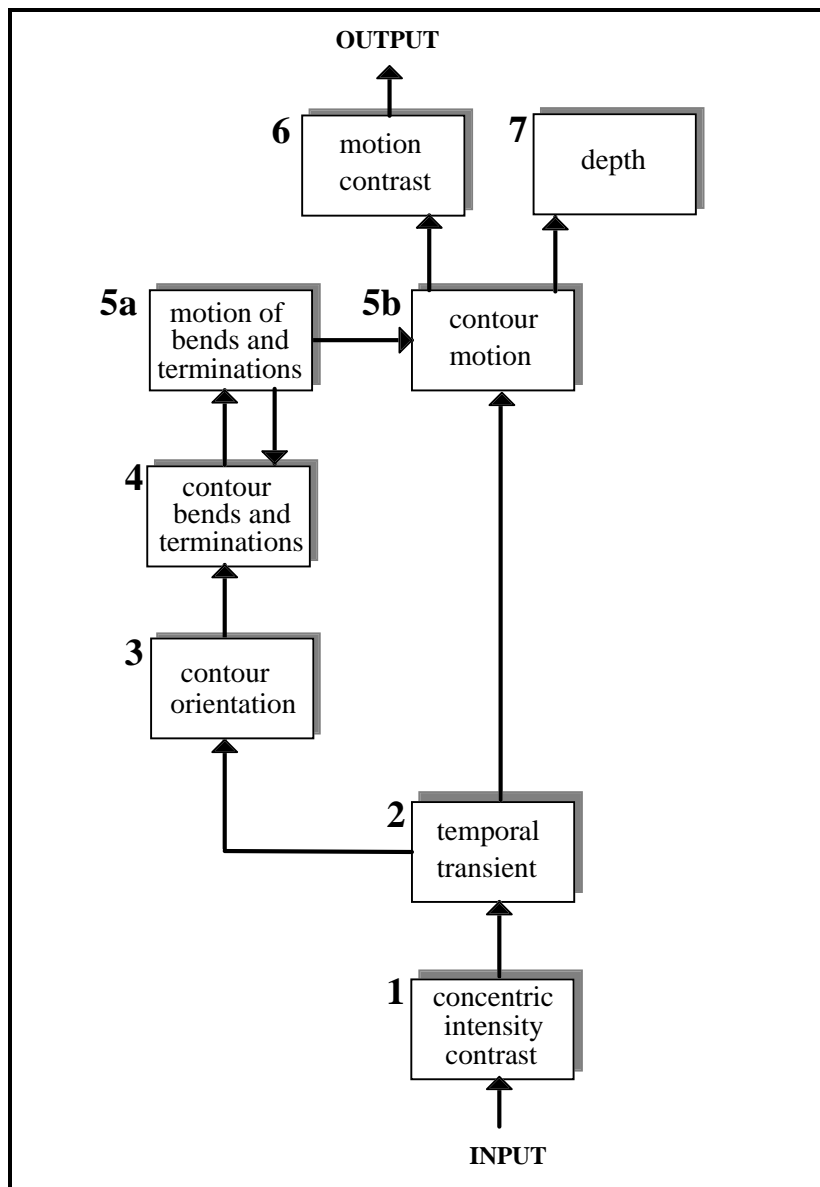
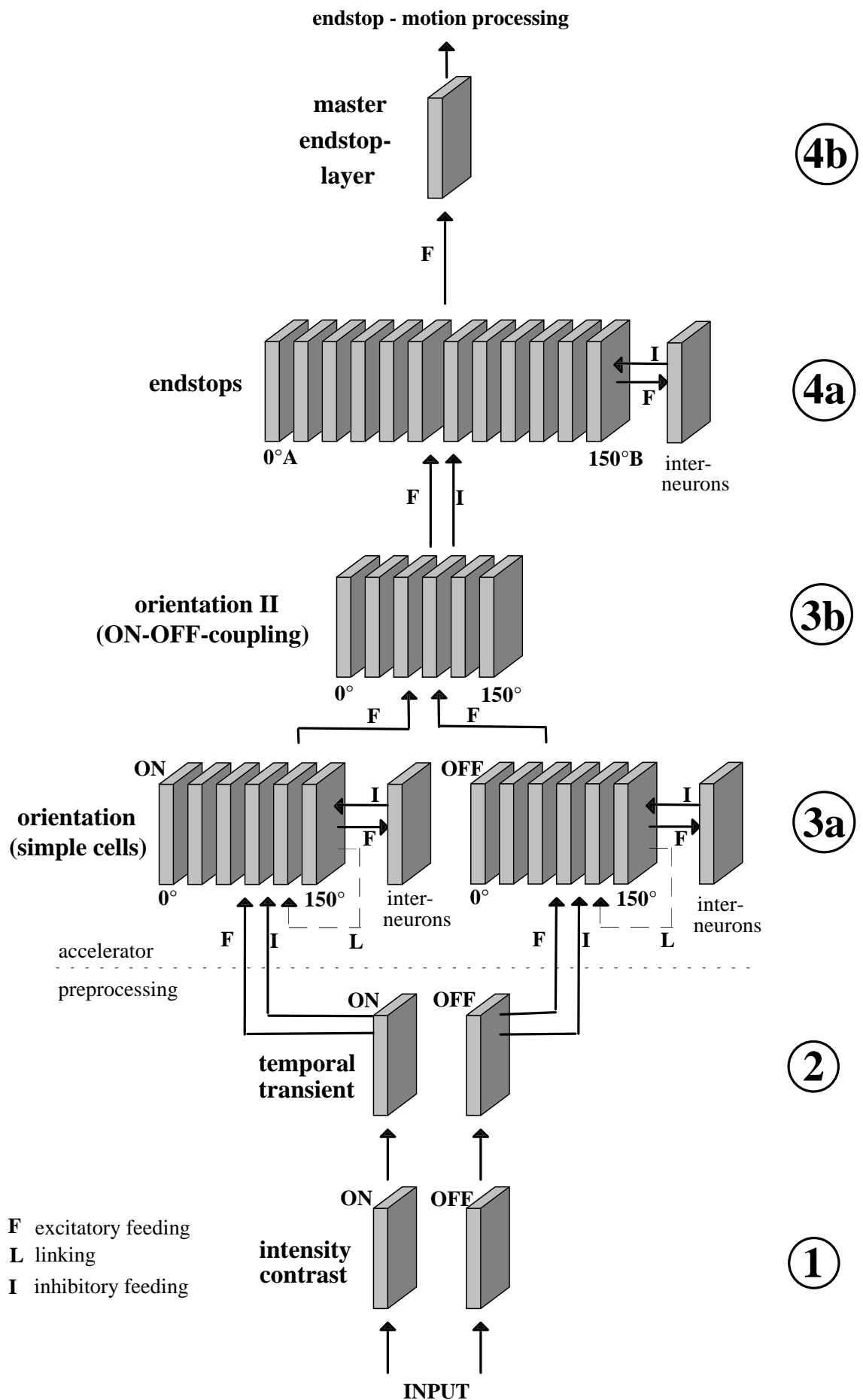


Abb. 3.77: Netzwerkmodell Transientensystem.

Dargestellt sind die einzelnen Verarbeitungsstufen von der visuellen Eingabe bis hin zur Objektdefinition durch Bewegungskontrast. Die Verarbeitungsstufen 1 und 2 befinden sich in der Vorverarbeitung, alle anderen auf dem Accelerator. **1:** Intensitätskontrastdetektion durch konzentrische receptive Felder mit Zentrum/Umfeld-Antagonismus. **2:** Detektion zeitlicher Änderungen der Intensitätskontrastkanten. **3:** Extraktion zusammenhängender Intensitätskontrastkanten und deren Orientierung. **4:** Bestimmung der Linienenden. **5a:** Richtungsdetektion der Linienenden mit Voraktivierung der Linienendendetektoren zur Detektionsverbesserung. **5b:** Richtungsdetektion der Kontur. **6:** Extraktion des lokalen Bewegungskontrastes zur Objektdefinition und zur Ansteuerung des Aufmerksamkeitsfokus. **7:** Bestimmung der relativen Objektiefe durch Bewegungsparallaxe.

rechts: Abb. 3.78: Modell zur Extraktion der Linienenden. Die Extraktion der Linienenden erfolgt in mehreren Verarbeitungsschritten. Die Vorverarbeitung (1, 2) extrahiert zunächst in getrennten ON- und OFF-Pfaden sich zeitlich ändernde Intensitätskontraste (Y-Zellen). Anschließend werden Intensitätskontrastkanten gleicher Orientierung mit einer bestimmten Mindestlänge detektiert (3a). Laterale Linking-Verbindungen verbessern hierbei das Detektionsergebnis. Auf der nachfolgenden Stufe (3b) werden ON- und OFF-Pfad vereinigt. In 4a werden die Linienenden extrahiert und in 4b zur nachfolgenden Bewegungsverarbeitung gesammelt.



3.4.1.1 Modell zur Orientierungsdetektion

Nachfolgend wird zunächst die Orientierungsdetektion und danach die Extraktion der Linienenden dargestellt. Die Orientierungsdetektion mit Marburger Modellneuronen wurde bereits in Ostkamp (1996), Spengler et al. (1996), Weitzel et al. (1997a, b) und Weitzel (1998) beschrieben. Die in dieser Arbeit dargestellte Orientierungsdetektion unterscheidet sich aber von jenen in zwei grundlegenden Punkten: Zum einen erfolgt die Detektion orientierter Intensitätskontrastkanten mit Detektoren, die ein symmetrisches rezeptives Feld haben, zum anderen ist das zeitliche Übertragungsverhalten der Modellneuronen anders als in den aufgeführten Arbeiten. Beide Unterschiede folgen aus der hier vorliegenden dynamischen Netzwerkeingabe.

Die Orientierungsdetektoren erhalten ihre Eingabe von Modellneuronen mit zeitlichen Bandpaßeigenschaften, den Y-Zellen. Die Y-Zellen werden im sogenannten *Burst*-Modus (Dicke 1992, Rausch et al. 1993, Erb 1993) betrieben. Dadurch befinden sich die Orientierungsdetektoren ebenfalls im Burst-Modus und nicht im Oszillations-Modus. Die Orientierungsdetektion mit Modellneuronen, deren neuronale Aktivität oszilliert, ist in den oben aufgeführten Arbeiten beschrieben. In beiden Modi erweisen sich laterale Interaktionen zwischen den Detektoren über Linking-Verbindungen als vorteilhaft.

Durch die Implementierung symmetrischer rezeptiver Felder kann die Polarität der Kante nicht mehr bestimmt werden. Dies ist aber für die Extraktion der Linienenden auch nicht erforderlich. Vielmehr reduziert sich der Bedarf an Modellneuronen dadurch auf die Hälfte. Außerdem sind Orientierungsdetektoren mit asymmetrischem rezeptiven Feld zusätzlich richtungssensitiv. Überstreicht eine Intensitätskontrastkante zunächst den exzitatorischen Teil des rezeptiven Feldes, führt sie zu einer Erhöhung des Membranpotentials und damit zu einer Aktivierung des Orientierungsdetektors. Die Orientierung der Kante kann dabei sogar um $\pm 90^\circ$ von der Orientierungspräferenz des Modellneurons abweichen. Modellneuronen mit symmetrischem rezeptiven Feld werden hingegen nur aktiv, wenn die Kantenorientierung der Orientierungspräferenz ähnlich ist, da die Kante zuvor ein inhibitorisches Teilgebiet durchläuft und somit eine Hyperpolarisation des Modellneurons verursacht hat.

Im hexagonalen Abtastgitter bieten sich Orientierungsdetektoren mit einem Unterschied in der Orientierungspräferenz von 60° , 30° oder 15° an. Als Kompromiß zwischen Modellneuronenanzahl und Auflösung wurde ein Orientierungspräferenzunterschied von 30° gewählt. Damit steht an jedem Abtastpunkt jeweils ein Satz von 6 ON- und 6 OFF-Orientierungsdetektoren zur Verfügung. Die Abbildungen 3.79 und 3.80 zeigen die Verschaltung der Ganglienzellen zu Orientierungsdetektoren. Die Gewichte (Gl 3.55 - 3.58) wurden so gewählt, daß in der Mitte ein exzitatorischer Bereich entsteht, der von zwei inhibitorischen Feldern umgeben ist. Die Gewichte sind für ON- und OFF-Pfad gleich.

(Gl. 3.55)

$$\begin{aligned}
\text{feed2}^Y w_{ijkl}^{\text{OD},0^\circ} = & \left\{ \begin{array}{ll} 4,5 \cdot {}^{\text{OD}}V^{\text{F2}}, & \text{falls } \left((i = k) \wedge ((j = l) \vee (j = (l + d))) \right) \vee \\ & \left((i = (k - d)) \wedge ((j = (l \pm d / 2))) \right) \\ 3 \cdot {}^{\text{OD}}V^{\text{F2}}, & \text{falls } \left((i = k) \wedge (j = (l - d)) \right) \vee \left((i = (k - d)) \wedge (j = (l + 3 \cdot d / 2)) \right) \\ -2 \cdot {}^{\text{OD}}V^{\text{F2}}, & \text{falls } \left((i = (k - 3 \cdot d)) \wedge ((j = (l - 3 \cdot d / 2)) \vee (j = (l + 5 \cdot d / 2))) \right) \vee \\ & \left((i = (k + 2 \cdot d)) \wedge (j = \pm 2 \cdot d) \right) \\ -3 \cdot {}^{\text{OD}}V^{\text{F2}}, & \text{falls } \left((i = (k - 3 \cdot d)) \wedge ((j = (l \pm d / 2)) \vee (j = (l + 3 \cdot d / 2))) \right) \vee \\ & \left((i = (k + 2 \cdot d)) \wedge ((j = l) \vee (j = (l \pm d))) \right) \\ -5 \cdot {}^{\text{OD}}V^{\text{F2}}, & \text{falls } \left((i = (k - 2 \cdot d)) \wedge ((j = (l - d)) \vee (j = (l + 2 \cdot d))) \right) \vee \\ & \left((i = (k + d)) \wedge (j = (l \pm 3 \cdot d / 2)) \right) \\ -6 \cdot {}^{\text{OD}}V^{\text{F2}}, & \text{falls } \left((i = (k - 2 \cdot d)) \wedge ((j = l) \vee (j = (l + d))) \right) \vee \\ & \left((i = (k + d)) \wedge (j = (l \pm d / 2)) \right) \\ 0, & \text{sonst} \end{array} \right.
\end{aligned}$$

(Gl. 3.56)

$$\begin{aligned}
\text{feed2}^Y w_{ijkl}^{\text{OD},30^\circ} = & \left\{ \begin{array}{ll} 5,5 \cdot {}^{\text{OD}}V^{\text{F2}}, & \text{falls } \left((i = (k - d)) \wedge (j = (l + 3 \cdot d / 2)) \right) \vee \left((i = k) \wedge (j = l) \right) \\ 5 \cdot {}^{\text{OD}}V^{\text{F2}}, & \text{falls } \left((i = (k - d)) \wedge (j = (l + 5 \cdot d / 2)) \right) \vee \left((i = k) \wedge (j = (l - d)) \right) \\ 4,5 \cdot {}^{\text{OD}}V^{\text{F2}}, & \text{falls } \left((i = (k - d)) \wedge (j = (l + d / 2)) \right) \vee \left((i = k) \wedge (j = (l + d / 2)) \right) \\ -3,5 \cdot {}^{\text{OD}}V^{\text{F2}}, & \text{falls } \left((i = (k \pm d)) \wedge (j = (l - d / 2)) \right) \vee \\ & \left(((i = k) \vee (i = (k - 2 \cdot d))) \wedge (j = (l + 2 \cdot d)) \right) \\ -4 \cdot {}^{\text{OD}}V^{\text{F2}}, & \text{falls } \left((i = k) \wedge ((j = (l - 2 \cdot d)) \vee (j = (l + 4 \cdot d))) \right) \vee \\ & \left((i = (k - d)) \wedge ((j = (l - 5 \cdot d / 2)) \vee (j = (l + 7 \cdot d / 2))) \right) \\ -5 \cdot {}^{\text{OD}}V^{\text{F2}}, & \text{falls } \left((i = (k - 2 \cdot d)) \wedge ((j = l) \vee (j = (l \pm d))) \right) \vee \\ & \left((i = (k - d)) \wedge (j = (l - 3 \cdot d / 2)) \right) \vee \\ & \left((i = k) \wedge (j = (l + 3 \cdot d)) \right) \vee \\ & \left((i = (k + d)) \wedge \right. \\ & \left. \left((j = (l + d / 2)) \vee (j = (l + 3 \cdot d / 2)) \vee (j = (l + 5 \cdot d / 2)) \right) \right) \\ 0, & \text{sonst} \end{array} \right.
\end{aligned}$$

(Gl. 3.57)

$$\begin{aligned}
{}^Y_{\text{feed2}} W_{ijkl}^{\text{OD}, 60^\circ} = & \left\{ \begin{array}{ll} 4 \cdot {}^{\text{OD}}V^{\text{F2}}, & \text{falls } \left((i = k) \wedge ((j = l) \vee (j = (l - d))) \right) \vee \\ & \left((i = (k - d)) \wedge ((j = (l \pm d / 2))) \right) \vee \\ & \left((i = (k + d)) \wedge ((j = (l - d / 2)) \vee (j = (l - 3 \cdot d / 2))) \right) \\ \\ -6 \cdot {}^{\text{OD}}V^{\text{F2}}, & \text{falls } \left((i = k) \wedge ((j = (l + d)) \vee (j = (l - 2 \cdot d))) \right) \vee \\ & \left((i = (k - d)) \wedge (j = (l \pm 3 \cdot d / 2)) \right) \vee \\ & \left((i = (k + d)) \wedge ((j = (l + d / 2)) \vee (j = (l - 5 \cdot d / 2))) \right) \\ \\ -3 \cdot {}^{\text{OD}}V^{\text{F2}}, & \text{falls } \left((i = k) \wedge ((j = (l + 2 \cdot d)) \vee (j = (l - 3 \cdot d))) \right) \vee \\ & \left((i = (k - d)) \wedge (j = (l \pm 5 \cdot d / 2)) \right) \vee \\ & \left((i = (k + d)) \wedge ((j = (l + 3 \cdot d / 2)) \vee (j = (l - 7 \cdot d / 2))) \right) \\ \\ -4,5 \cdot {}^{\text{OD}}V^{\text{F2}}, & \text{falls } \left((i = (k + 2 \cdot d)) \wedge (j = l) \right) \vee \left((i = (k - 2 \cdot d)) \wedge (j = (l - d)) \right) \\ \\ -1,5 \cdot {}^{\text{OD}}V^{\text{F2}}, & \text{falls } \left((i = (k + 2 \cdot d)) \wedge (j = (l - 3 \cdot d)) \right) \vee \\ & \left((i = (k - 2 \cdot d)) \wedge (j = (l + 2 \cdot d)) \right) \\ \\ 0, & \text{sonst} \end{array} \right.
\end{aligned}$$

(Gl. 3.58)

$$\begin{aligned}
{}^Y_{\text{feed2}} W_{ijkl}^{\text{OD}, 90^\circ} = & \left\{ \begin{array}{ll} 6 \cdot {}^{\text{OD}}V^{\text{F2}}, & \text{falls } \left((i = k) \wedge ((j = l) \vee (j = (l - d))) \right) \vee \\ & \left((i = (k \pm d)) \wedge (j = (l - d / 2)) \right) \\ \\ -6 \cdot {}^{\text{OD}}V^{\text{F2}}, & \text{falls } \left((i = k) \wedge ((j = (l - 2 \cdot d)) \vee (j = (l + d))) \right) \vee \\ & \left((i = (k \pm d)) \wedge ((j = (l - 3 \cdot d / 2)) \vee (j = (l + d / 2))) \right) \\ \\ -3 \cdot {}^{\text{OD}}V^{\text{F2}}, & \text{falls } \left((i = (k \pm d)) \wedge ((j = (l - 5 \cdot d / 2)) \vee (j = (l + 3 \cdot d / 2))) \right) \vee \\ & \left((i = k) \wedge ((j = (l - 3 \cdot d)) \vee (j = (l + 2 \cdot d))) \right) \vee \\ & \left((i = (k \pm 2 \cdot d)) \wedge ((j = l) \vee (j = (l - d))) \right) \\ \\ 0, & \text{sonst} \end{array} \right.
\end{aligned}$$

Die Zeitkonstanten der Feeding-Eingänge wurden gleich groß gewählt. Die Zeitkonstanten dürfen nicht zu groß sein, damit sich bei bewegten Objekten keine „Aktivitätsschweife“ bilden. Da der Orientierungsdetektor weiterhin Klassifikatoreigenschaften haben soll, müssen aufgrund des Verhältnisses zwischen Eingangsgewichten und Schwellenoffset die einzelnen Teilpotentiale möglichst schnell auf ihren Ausgangswert abklingen (Tabelle 3.8).

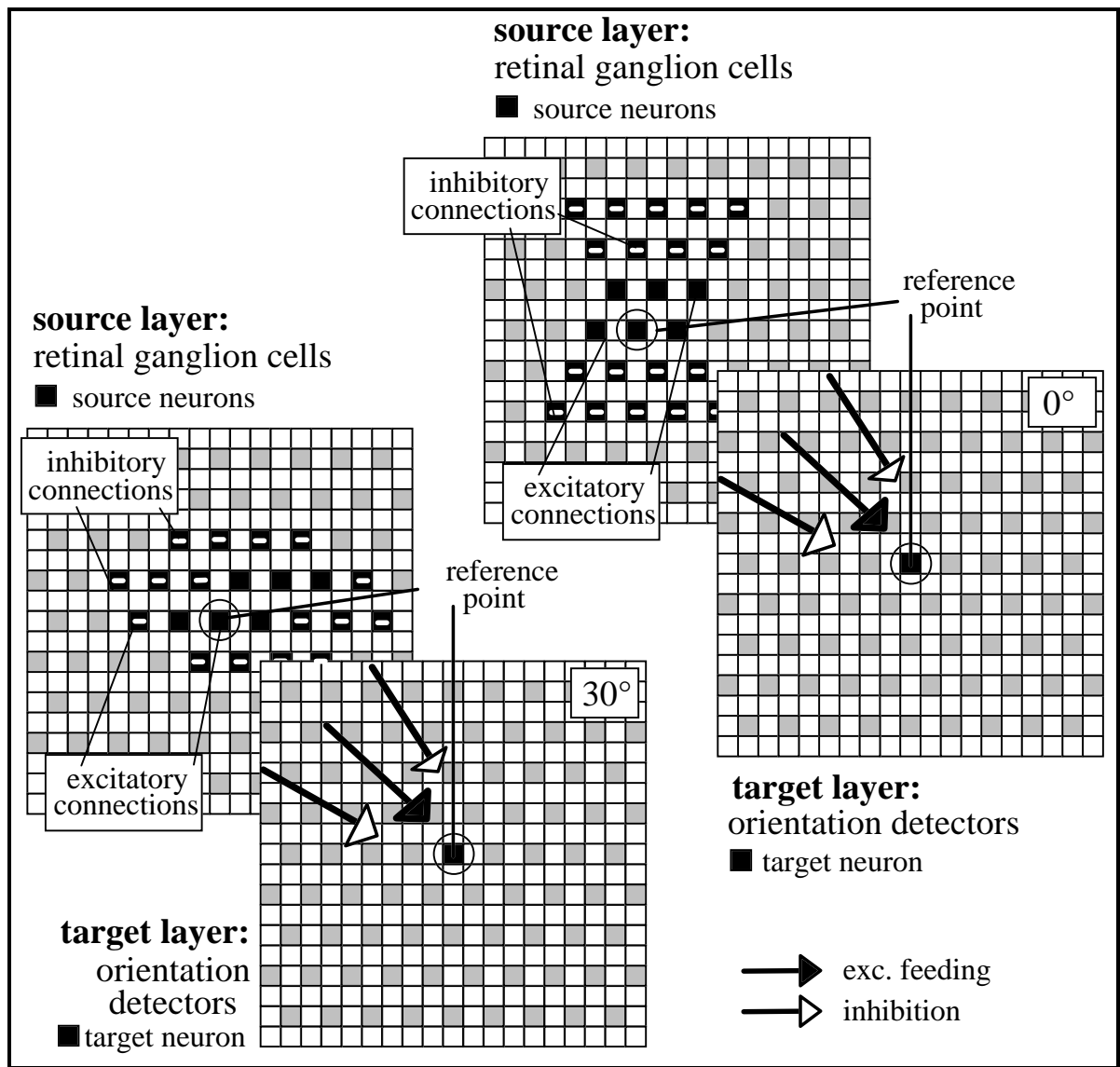


Abb. 3.79: Verschaltung der Ganglienzellen zu 0°- und 30°-Orientierungsdetektoren. Dargestellt sind die Verschaltungen der retinalen Ganglienzellen (schwarze Rechtecke) zu Detektoren mit einer Orientierungspräferenz von 0° und 30°. Die Verschaltung der 150°-Orientierungsdetektoren ergibt sich durch Spiegelung der 30°-Orientierungsdetektoren an der vertikalen Achse.

Zur Verbesserung der Detektionseigenschaften wurden wie auch bei der Bewegungsdetektion laterale Linking-Verbindungen (Abbildung 3.81, Gleichungen 3.59 bis 3.62) und inhibitorisch wirkende Interneuronen implementiert. Dieser Verschaltung liegt das Gestaltprinzip des guten Verlaufs zugrunde: Der Verlauf einer Kante ändert sich in den meisten Fällen nur gering. Demzufolge ist eine laterale Linking-Verschaltung zwischen benachbarten Orientierungsdetektoren gleicher Orientierungspräferenz in Richtung ihrer Orientierungspräferenz sinnvoll. Damit auch Kanten mit kleinen Orientierungsabweichungen zur Vorzugsorientierung der Detektoren sicher detektiert werden können, werden zusätzlich Linking-Verbindungen in Richtungen ähnlich der Orientierungspräferenz implementiert, wodurch sich eine keulenförmige Verschaltungsstruktur der Linking-Verbindungen ergibt. Zwischen ON- und OFF-Pfad gibt es in diesem Modell keine Linking-Verbindungen. Die angegebenen Gewichte sind für ON- und OFF-Pfad gültig.

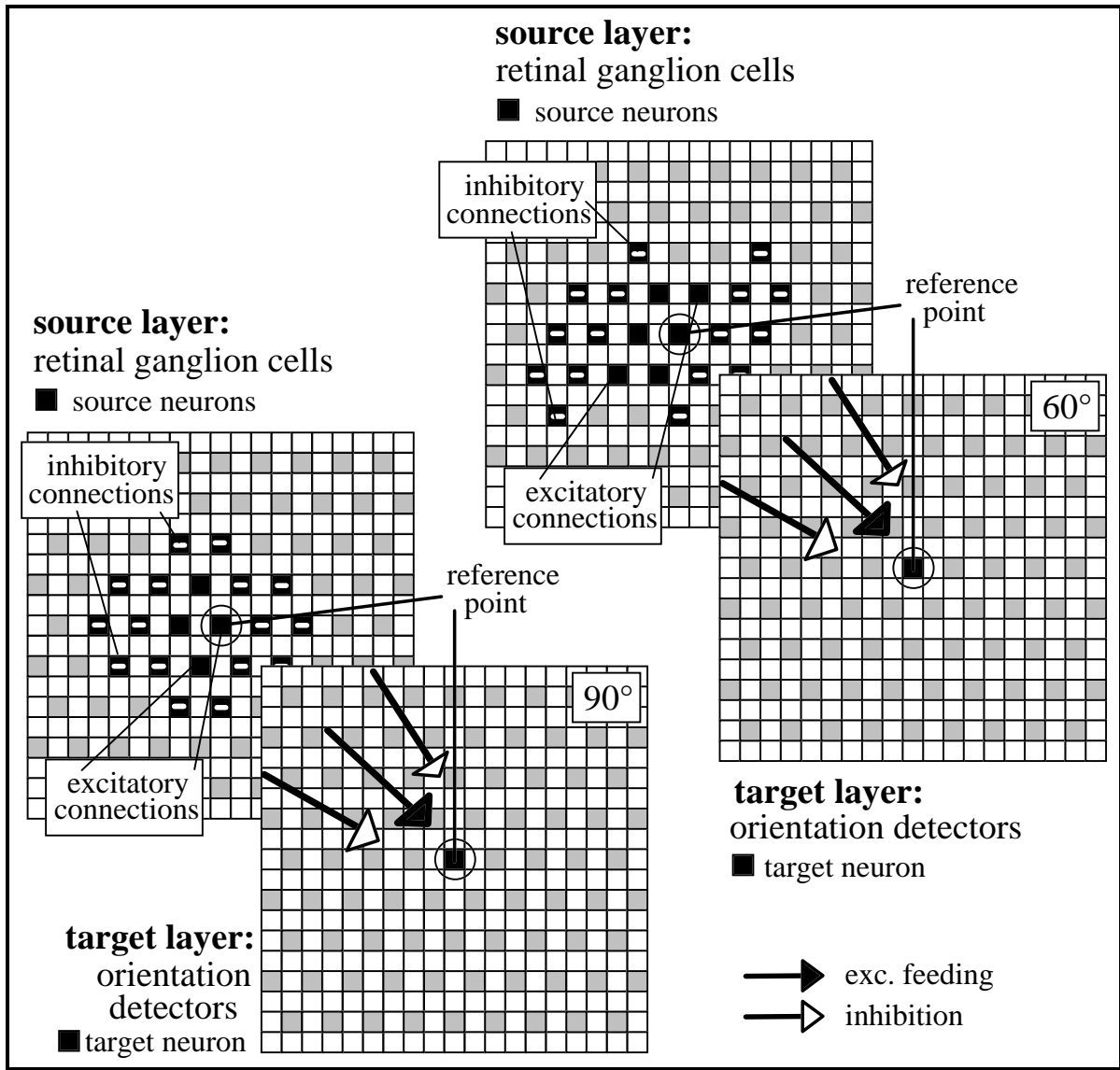


Abb. 3.80: Verschaltung der Ganglienzellen zu 60°- und 90°-Orientierungsdetektoren. Dargestellt sind die Verschaltungen der retinalen Ganglienzellen (schwarze Rechtecke) zu Detektoren mit einer Orientierungspräferenz von 60° und 90°. Die Verschaltung der 120°-Orientierungsdetektoren ergeben sich durch Spiegelung der 60°-Orientierungsdetektoren an der vertikalen Achse.

(Gl. 3.59)

$${}_{\text{link}}^{OD,0^\circ} w_{ijkl}^{OD,0^\circ} = \begin{cases} 0,25 \cdot {}^{OD}V^L, & \text{falls } ((i = k) \wedge (j = (l \pm d))) \\ 0,0625 \cdot {}^{OD}V^L, & \text{falls } ((i = k) \wedge (j = (l \pm 2 \cdot d))) \\ 0,125 \cdot {}^{OD}V^L, & \begin{aligned} &\text{falls } (i = (k \pm d)) \wedge \\ &((j = (l \pm 3 \cdot d / 2)) \vee (j = (l \pm 5 \cdot d / 2)) \vee (j = (l \pm 7 \cdot d / 2))) \end{aligned} \\ 0, & \text{sonst} \end{cases}$$

(Gl. 3.60)

$${}^{\text{OD},30^\circ}_{\text{link}} w_{ijkl}^{\text{OD},30^\circ} = \begin{cases} 0,125 \cdot {}^{\text{OD}}V^L, & \text{falls } \left((i = (k+d)) \wedge ((j = (l-5 \cdot d/2)) \vee (j = (l-7 \cdot d/2))) \right) \vee \\ & \left((i = (k+2 \cdot d)) \wedge ((j = (l-2 \cdot d)) \vee (j = (l-4 \cdot d))) \right) \vee \\ & \left((i = (k+3 \cdot d)) \wedge ((j = (l-5 \cdot d/2)) \vee (j = (l-7 \cdot d/2))) \right) \vee \\ & \left((i = (k-d)) \wedge ((j = (l+5 \cdot d/2)) \vee (j = (l+7 \cdot d/2))) \right) \vee \\ & \left((i = (k-2 \cdot d)) \wedge ((j = (l+2 \cdot d)) \vee (j = (l+4 \cdot d))) \right) \vee \\ & \left((i = (k-3 \cdot d)) \wedge ((j = (l+5 \cdot d/2)) \vee (j = (l+7 \cdot d/2))) \right) \\ 0,0625 \cdot {}^{\text{OD}}V^L, & \text{falls } \left((i = (k+2 \cdot d)) \wedge (j = (l-3 \cdot d)) \right) \vee \\ & \left((i = (k-2 \cdot d)) \wedge (j = (l+3 \cdot d)) \right) \\ 0,25 \cdot {}^{\text{OD}}V^L, & \text{falls } \left((i = (k+d)) \wedge (j = (l-3 \cdot d/2)) \right) \vee \\ & \left((i = (k-d)) \wedge (j = (l+3 \cdot d/2)) \right) \\ 0, & \text{sonst} \end{cases}$$

(Gl. 3.61)

$${}^{\text{OD},30^\circ}_{\text{link}} w_{ijkl}^{\text{OD},30^\circ} = \begin{cases} 0,125 \cdot {}^{\text{OD}}V^L, & \text{falls } \left((i = (k-d)) \wedge ((j = (l+3 \cdot d/2))) \right) \vee \\ & \left((i = (k-2 \cdot d)) \wedge ((j = l) \vee (j = (l+2 \cdot d))) \right) \vee \\ & \left((i = (k-3 \cdot d)) \wedge ((j = (l+d/2)) \vee (j = (l+5 \cdot d/2))) \right) \vee \\ & \left((i = (k-4 \cdot d)) \wedge (j = (l+d)) \right) \vee \left((i = (k+d)) \wedge (j = (l-3 \cdot d/2)) \right) \\ & \left((i = (k+3 \cdot d)) \wedge ((j = (l-d/2)) \vee (j = (l-5 \cdot d/2))) \right) \vee \\ & \left((i = (k+2 \cdot d)) \wedge ((j = l) \vee (j = (l-2 \cdot d))) \right) \vee \\ & \left((i = (k+4 \cdot d)) \wedge (j = (l-d)) \right) \\ 0,0625 \cdot {}^{\text{OD}}V^L, & \text{falls } \left((i = (k-2 \cdot d)) \wedge (j = (l+d)) \right) \vee \\ & \left((i = (k+2 \cdot d)) \wedge (j = (l-d)) \right) \\ 0,25 \cdot {}^{\text{OD}}V^L, & \text{falls } \left((i = (k-d)) \wedge (j = (l+d/2)) \right) \vee \\ & \left((i = (k+d)) \wedge (j = (l-d/2)) \right) \\ 0, & \text{sonst} \end{cases}$$

(Gl. 3.62)

$${}^{\text{OD},90^\circ}_{\text{link}} w_{ijkl}^{\text{OD},90^\circ} = \begin{cases} 0,4375 \cdot {}^{\text{OD}}V^L & \text{falls } \left((i = (k \pm 2 \cdot d)) \wedge (j = l) \right) \vee \\ & \left((i = (k \pm 3 \cdot d)) \wedge (j = (l \pm d/2)) \right) \\ 0,1875 \cdot {}^{\text{OD}}V^L & \text{falls } \left((i = (k \pm 4 \cdot d)) \wedge (j = (l \pm d)) \right) \vee \\ & \left((i = (k \pm 5 \cdot d)) \wedge (j = (l \pm d/2)) \right) \\ 0,125 \cdot {}^{\text{OD}}V^L & \text{falls } \left((i = (k \pm 4)) \wedge (j = l) \right) \\ 0, & \text{sonst} \end{cases}$$

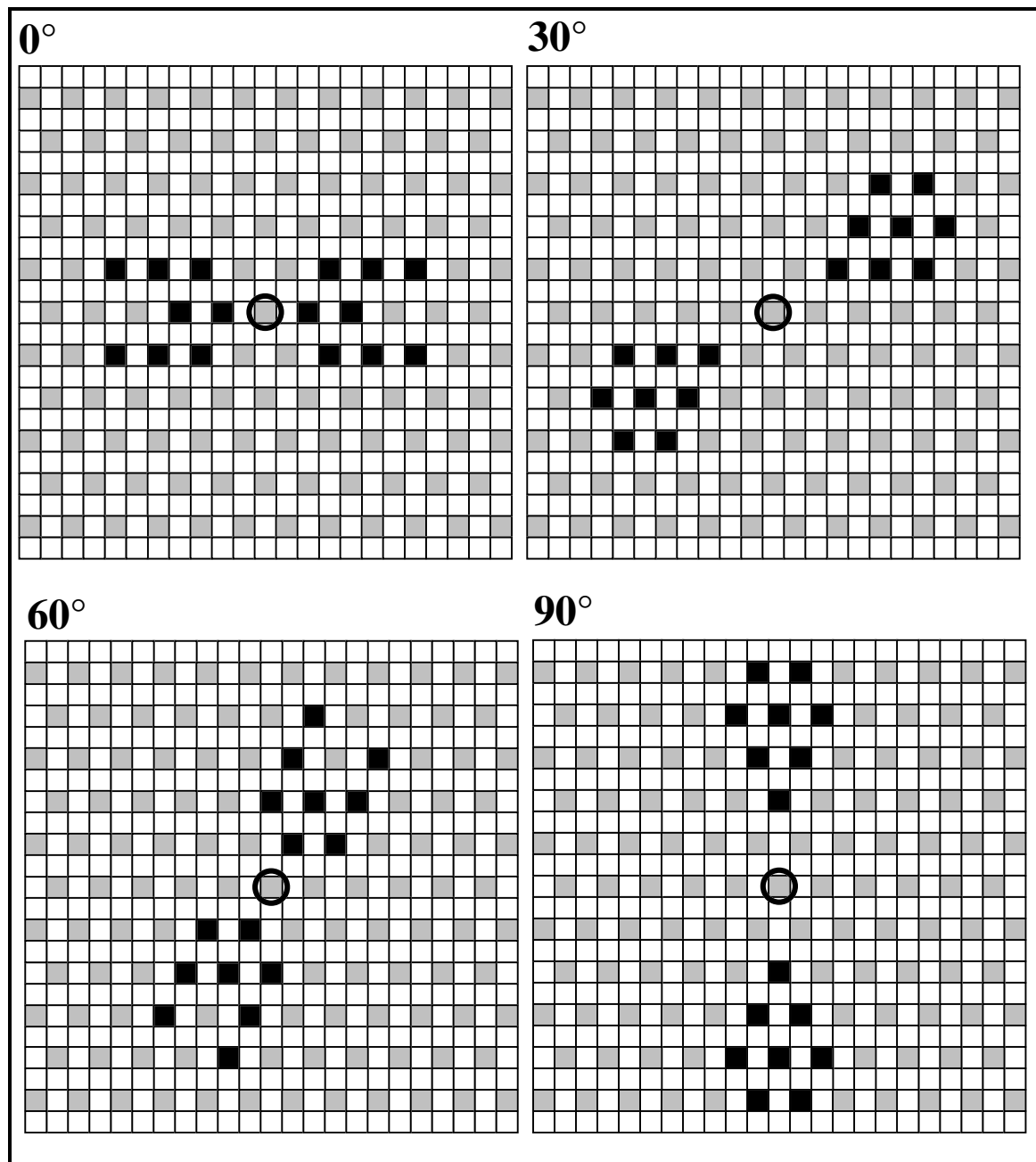


Abb. 3.81: Linking-Verbindungen der Orientierungsdetektoren. Die Abbildung zeigt die Linking-Verbindungen der Orientierungsdetektoren innerhalb einer Orientierungsdetektorschicht anhand jeweils eines ausgewählten Orientierungsdetektors (Kreis). Die Orientierungsdetektoren, deren RF-Mittelpunkte schwarz markiert sind, erhalten laterale Linking-Eingabe von dem Orientierungsdetektor des Bezugspunktes. Die Linking-Verschaltungen für Orientierungsdetektoren mit Kantenorientierungspräferenzen von 120° und 150° ergeben sich aus Spiegelung der Linking-Verschaltung der 60°- bzw. 30°-Orientierungsdetektoren an der vertikalen Achse.

Zur Beschränkung der Aktivität der Orientierungsdetektoren und zur Verschärfung der Orientierungscharakteristik sind zusätzliche Inhibitionsneuronen implementiert. Für jeden Orientierungsdetektorsatz existiert ein Inhibitionsneuron. Dieses wird aufgrund seines hohen Schwellenoffsets erst bei hinreichend großer Aktivität der Modellneuronen eines Orientierungsdetektorsatzes aktiv. Das Interneuron hemmt diesen Orientierungsdetektorsatz und seine sechs räumlich benachbarten Orientierungsdetektorsätze, was unter Beachtung der Linking-Wechselwirkungen zu einer effektiven Hemmung aller Orientierungsdetektoren mit Aus-

nahme des Detektors mit der am besten zur Kante passenden Orientierungspräferenz und Lage des rezeptiven Feldes führt.

Das Prinzip der Kombination aus lateraler Linking-Verschaltung und Inhibitionsneuron zur Verbesserung der Detektionseigenschaften ist bei der Orientierungsdetektion und der Bewegungsdetektion identisch. Unter Verwendung anderer Parameter, insbesondere größerer Zeitkonstanten, kann die dargestellte Methode auch im Oszillations-Modus erfolgreich eingesetzt werden (Spengler et al. 1996, Spengler 1996, Weitzel 1998).

Durch die Aufteilung der Konturrepräsentation in Schichten mit Detektoren verschiedener Orientierungspräferenz lassen sich Linienenden, Ecken und starke Krümmungen bestimmen. Bevor diese von den Linienendendetektoren extrahiert werden können, werden die Orientierungsdetektorergebnisse des ON- und des OFF-Pfades zusammengelegt. Dies ist vor allem eine Maßnahme, um die Anzahl der benötigten Linienendendetektoren zu reduzieren. Der Detektor (i, j) des ON-Pfades mit der Orientierungspräferenz ϕ' wird mit dem Detektor (k, l) der nachfolgenden Schicht OD-II über eine Feeding-Verbindung gemäß Gleichung 3.63 gekoppelt. Für den OFF-Pfad gilt entsprechend Gleichung 3.64. Die Modellneuronen der Schichten OD-II übernehmen somit die Orientierungspräferenz der vorgeschalteten Detektoren.

$$\text{ON-OD, } \phi'=\phi \text{ }_{\text{feed2}} w_{ijkl}^{\text{OD-II, } \phi} = \begin{cases} \text{OD-II } V^{F2}, & \text{falls } (i == k) \wedge (j == l) \\ 0, & \text{sonst} \end{cases} \quad (\text{Gl. 3.63})$$

$$\text{OFF-OD, } \phi'=\phi \text{ }_{\text{feed2}} w_{ijkl}^{\text{OD-II, } \phi} = \begin{cases} \text{OD-II } V^{F2}, & \text{falls } (i == k) \wedge (j == l) \\ 0, & \text{sonst} \end{cases} \quad (\text{Gl. 3.64})$$

mit $\phi \in \{0^\circ, 30^\circ, 60^\circ, 90^\circ, 120^\circ, 150^\circ\}$

Die Orientierungsdetektoren-II sind wiederum über Linking-Verbindungen gekoppelt, wie sie bereits in den Gleichungen 3.59 bis 3.62 beschrieben wurden. Die zugehörigen inhibitorisch wirkenden Interneuronen werden gemäß den Gleichungen 3.65 und 3.66 verschaltet.

$$\text{OD-II, } \phi' \text{ }_{\text{feed1}} w_{ijkl}^{\text{OD-II, Inter}} = \begin{cases} \text{IOD-II } V^{F1}, & \text{falls } (i == k) \wedge (j == l) \\ 0, & \text{sonst} \end{cases} \quad (\text{Gl. 3.65})$$

$$\text{OD-II, Inter }_{\text{inh}} w_{ijkl}^{\text{OD-II, } \phi} = \begin{cases} \text{OD-II } V^I, & \text{falls } ((i == k) \wedge (j == l)) \vee \text{hexagon}_d 1 \\ 0, & \text{sonst} \end{cases} \quad (\text{Gl. 3.66})$$

mit ϕ' und $\phi \in \{0^\circ, 30^\circ, 60^\circ, 90^\circ, 120^\circ, 150^\circ\}$

3.4.1.2 Modell zur Linienendendetektion

Für die Extraktion der Linienenden werden die in den Abbildungen 3.82 und 3.83 schematisch dargestellten Detektoren verwendet. Die Linienendendetektoren bekommen ihre Eingabe von Orientierungsdetektoren. Dies können entweder ausschließlich Orientierungsdetektoren des ON- oder OFF-Pfades sein oder die der zwischengeschalteten Schichten bei Kopplung von ON- und OFF-Pfad. Die Entscheidung, welche Orientierungsdetektoren auf die Linienendendetektoren verschaltet werden, hängt von den im technischen System verfügbaren Modellneuronen ab.

Die Verschaltungsstruktur eines Linienendendetektors besteht aus einem exzitatorisch und einem inhibitorisch wirkenden Teilgebiet (Abbildungen 3.82 und 3.83). Wird das Modell-

(Gl. 3.69)

$$OD-II,60^\circ_{feed2} W_{ijkl}^{ESD,60^\circ A} = \begin{cases} 5 \cdot^{ESD} V^{F2}, & \text{falls } ((i == k) \wedge (j == l)) \vee ((i == (k + d)) \wedge (j == (l - d / 2))) \\ & \vee ((i == (k + 2 \cdot d)) \wedge (j == (l - d))) \\ 0, & \text{sonst} \end{cases}$$

(Gl. 3.70)

$$OD-II,90^\circ_{feed2} W_{ijkl}^{ESD,90^\circ A} = \begin{cases} 4 \cdot^{ESD} V^{F2}, & \text{falls } ((i == k) \vee (i == (k + 2 \cdot d))) \wedge (j == l) \\ 0, & \text{sonst} \end{cases}$$

(Gl. 3.71)

$$OD-II,0^\circ_{inh} W_{ijkl}^{ESD,0^\circ A} = \begin{cases} ^{ESD} V^I, & \text{falls } ((i == k) \wedge ((j == (l + d)) \vee (j == (l + 2 \cdot d)) \vee (j == (l + 3 \cdot d)))) \\ & \vee \left((i == (k \pm d)) \wedge \left((j == (l + 3 \cdot d / 2)) \vee \right. \right. \\ & \left. \left. (j == (l + 5 \cdot d / 2)) \vee (j == (l + 7 \cdot d / 2)) \right) \right) \\ 0, & \text{sonst} \end{cases}$$

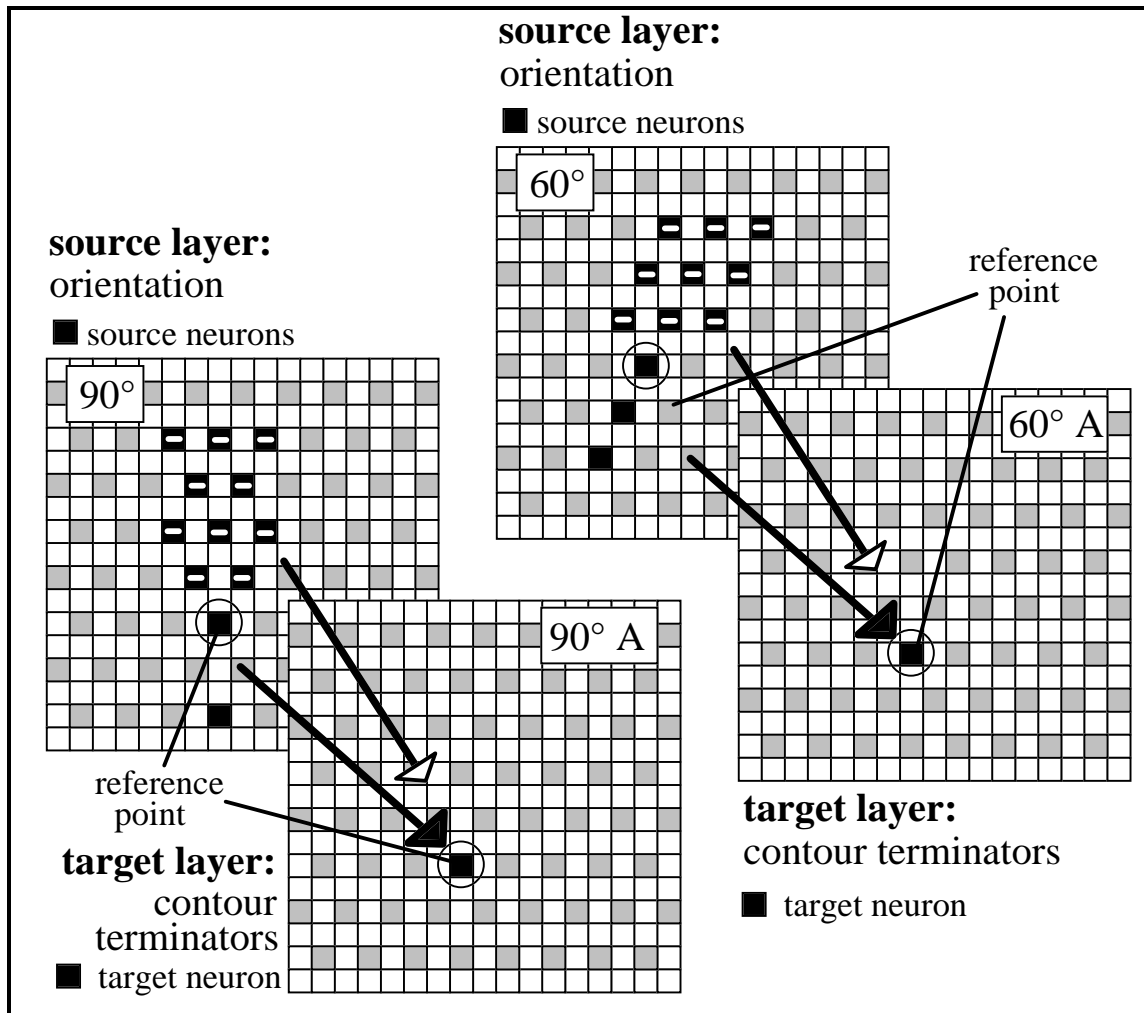


Abb. 3.83: Verschaltung der 60°- und 90°-Linienendendetektoren. Die Abbildung zeigt die Verschaltungen der 60°- und 90°-Orientierungsdetektoren (schwarze Rechtecke) zu Linienendendetektoren. Zu jeder Orientierung existieren zwei Linienendendetektoren. Durch eine 180°-Drehung um den Bezugspunkt (Kreis) ergeben sich die anderen Linienendendetektoren. Aus Gründen der Übersichtlichkeit wird nur jeweils die Verschaltung eines Linienendendetektors angegeben (Kreis).

(Gl. 3.72)

$$OD-II,30^\circ_{inh} W_{ijkl}^{ESD,30^\circ A} = \begin{cases} {}^{ESD}V^I, & \text{falls } \left((i == (k-d)) \wedge \left((j == (l+d/2)) \vee \right. \right. \\ & \left. \left. (j == (l+3 \cdot d/2)) \vee (j == (l+5 \cdot d/2)) \right) \right) \vee \\ & \left((i == (k-2 \cdot d)) \wedge \left((j == (l+2 \cdot d)) \vee \right. \right. \\ & \left. \left. (j == (l+3 \cdot d)) \vee (j == (l+4 \cdot d)) \right) \right) \vee \\ & \left((i == (k-3 \cdot d)) \wedge \left((j == (l+7 \cdot d/2)) \vee \right. \right. \\ & \left. \left. (j == (l+9 \cdot d/2)) \vee (j == (l+11 \cdot d/2)) \right) \right) \\ 0, & \text{sonst} \end{cases}$$

(Gl. 3.73)

$$OD-II,60^\circ_{inh} W_{ijkl}^{ESD,60^\circ A} = \begin{cases} {}^{ESD}V^I, & \text{falls } \left((i == (k-d)) \wedge \left((j == (l \pm d/2)) \vee (j == (l+3 \cdot d/2)) \right) \right) \vee \\ & \left((i == (k-2 \cdot d)) \wedge \left((j == l) \vee (j == (l+d)) \vee (j == (l+2 \cdot d)) \right) \right) \vee \\ & \left((i == (k-3 \cdot d)) \wedge \left((j == (l+d/2)) \vee \right. \right. \\ & \left. \left. (j == (l+3 \cdot d/2)) \vee (j == (l+5 \cdot d/2)) \right) \right) \\ 0, & \text{sonst} \end{cases}$$

(Gl. 3.74)

$$OD-II,90^\circ_{inh} W_{ijkl}^{ESD,90^\circ A} = \begin{cases} {}^{ESD}V^I, & \text{falls } \left(\left((i == (k-d)) \vee (i == (k-3 \cdot d)) \right) \wedge (j == (l \pm d/2)) \right) \vee \\ & \left(\left((i == (k-2 \cdot d)) \vee (i == (k-4 \cdot d)) \right) \wedge \left((j == l) \vee (j == (l \pm d)) \right) \right) \\ 0, & \text{sonst} \end{cases}$$

Der große Vorteil bei der Bewegungsrichtungsbestimmung der Linienenden ist das geringe Korrespondenzproblem, da es im Vergleich zur gesamten Kontur nur wenige Linienenden im Eingabebild gibt. Damit dieser Vorteil auch ausgenutzt werden kann, darf eine Ecke oder ein Linienende im Eingabebild auch nur genau an einer Stelle detektiert werden. Dies ist in der Regel nicht der Fall, da durch die Größe des exzitatorischen Teilgebietes immer mehrere Linienendendetektoren mit räumlich benachbarten rezeptiven Feldern eine Membranpotentialerhöhung erfahren. Dadurch wird es nötig, die Detektion über einen inhibitorischen Rückkopplungskreis eindeutig zu machen. Diese inhibitorische Rückkopplung wird in Analogie zu der der Bewegungsrichtungs- und Orientierungsdetektoren implementiert. Die Ausgaben der Linienendendetektoren werden gemäß Gleichung 3.75 auf Interneuronen gegeben. Diese hemmen wiederum die Linienendendetektoren nach Gleichung 3.76. Im Unterschied zu der inhibitorischen Rückkopplung der Bewegungsrichtungs- und Orientierungsdetektoren greift die der Linienendendetektoren weiter aus. Die Inhibition der Bewegungsrichtungs- und Orientierungsdetektoren hat die Aufgabe, die Eindeutigkeit der detektierten Richtung bzw. Orientierung zu gewährleisten, wohingegen die Inhibition der Linienenden primär eine räumlich eindeutige Detektion ermöglichen soll.

$$ESD,\phi'A/B_{feedI} W_{ijkl}^{ESD,Inter} = \begin{cases} {}^{IESD}V^{F1}, & \text{falls } ((i == k) \wedge (j == l)) \\ 0, & \text{sonst} \end{cases} \quad (Gl. 3.75)$$

$${}_{\text{inh}}^{\text{ESD,Inter}} w_{ijkl}^{\text{ESD},\phi A/B} = \begin{cases} {}^{\text{ESD}} V^{I'}, & \text{falls } ((i == k) \wedge (j == l)) \vee \text{hexagon}_d 1 \vee \text{hexagon}_d 2 \\ 0, & \text{sonst} \end{cases} \quad (\text{Gl. 3.76})$$

mit ϕ' und $\phi \in \{0^\circ, 30^\circ, 60^\circ, 90^\circ, 120^\circ, 150^\circ\}$

3.4.1.3 Modell zur Bestimmung der Linienendenbewegungsrichtung

Nach der Extraktion der Linienenden soll deren Bewegungsrichtung bestimmt werden. Das neuronale Netzwerk zur Bestimmung der Linienendenbewegungsrichtung ist in der Abbildung 3.84 dargestellt. Vorteilhaft wäre eine Bestimmung der Bewegungsrichtung getrennt nach Linienendenorientierungen, da in diesem Fall das Korrespondenzproblem am geringsten wäre. Dies hätte allerdings zur Folge, daß je Linienendendetektorschicht sechs Richtungsdetektorschichten notwendig wären. Da dies eine zu große Menge an Modellneuronen beanspruchen würde, werden die in den Linienendendetektorschichten repräsentierten Linienenden nach Gleichung 3.77 in einer einzelnen übergeordneten Schicht - der *Master-Schicht* - gesammelt. Damit die Bewegungsrichtung für endliche Geschwindigkeiten bestimmt werden kann, muß neben einer unverzögerten Eingabe auch noch eine verzögerte vorliegen. Deswegen werden zwei weitere neuronale Schichten benötigt, deren Modellneuronen die Aufgabe haben, die Ausgabe der Master-Schicht um jeweils 16 Bin zu verzögern (Gleichungen 3.78 und 3.79). Hätten die Modellneuronen zwei Ausgänge, einen unverzögerten und einen verzögerten, dann könnten diese beiden zusätzlichen Schichten entfallen.

$${}_{\text{feed1}}^{\text{ESD},\phi A/B} w_{ijkl}^{\text{ESD,Master}} = \begin{cases} {}^{\text{ESD,Master}} V^{F1}, & \text{falls } ((i == k) \wedge (j == l)) \\ 0, & \text{sonst} \end{cases} \quad (\text{Gl. 3.77})$$

mit $\phi \in \{0^\circ, 30^\circ, 60^\circ, 90^\circ, 120^\circ, 150^\circ\}$

$${}_{\text{feed1}}^{\text{ESD,MasterDa}} w_{ijkl}^{\text{ESD,MasterDb}} = \begin{cases} {}^{\text{ESD,Master}} V^{F1}, & \text{falls } ((i == k) \wedge (j == l)) \\ 0, & \text{sonst} \end{cases} \quad (\text{Gl. 3.78})$$

$${}_{\text{feed1}}^{\text{ESD,MasterDa}} w_{ijkl}^{\text{ESD,MasterDb}} = \begin{cases} {}^{\text{ESD,Master}} V^{F1}, & \text{falls } ((i == k) \wedge (j == l)) \\ 0, & \text{sonst} \end{cases} \quad (\text{Gl. 3.79})$$

Die Richtungsdetektoren der Linienenden werden analog zu denen der Kontur, wie in Kapitel 3.1 dargestellt, konstruiert. Als Beispiel ist die Verschaltung des 0° -LEBRD in den Gleichungen 3.80 und 3.81 beschrieben. Im Unterschied zu den Richtungsdetektoren für Konturen muß hier aufgrund der geringeren Aktivität der Linienendendetektoren ein größerer Verstärkungsfaktor ${}^{\text{M-ESD}} V^{F1}$ benutzt werden. Der Parameter d' bezeichnet die Entfernung zwischen den beiden Abtastorten des bilokalen Richtungsdetektors und kann ein Vielfaches von d sein. Bei einer ausreichend großen Anzahl an Modellneuronen sollten LEBRD-Schichten mit verschiedenen d' implementiert werden, damit das neuronale Netzwerk für alle möglichen Linienendengeschwindigkeiten sensitiv ist.

$${}_{\text{feed1}}^{\text{ESD,Master}} w_{ijkl}^{\text{M-ESD},0^\circ} = \begin{cases} {}^{\text{M-ESD}} V^{F1}, & \text{falls } ((i == k) \wedge (j == l)) \\ 0, & \text{sonst} \end{cases} \quad (\text{Gl. 3.80})$$

$${}_{\text{feed1}}^{\text{ESD,MasterDb}} w_{ijkl}^{\text{M-ESD},0^\circ} = \begin{cases} {}^{\text{M-ESD}} V^{F1}, & \text{falls } ((i == k) \wedge (j == (l - d'))) \\ 0, & \text{sonst} \end{cases} \quad (\text{Gl. 3.81})$$

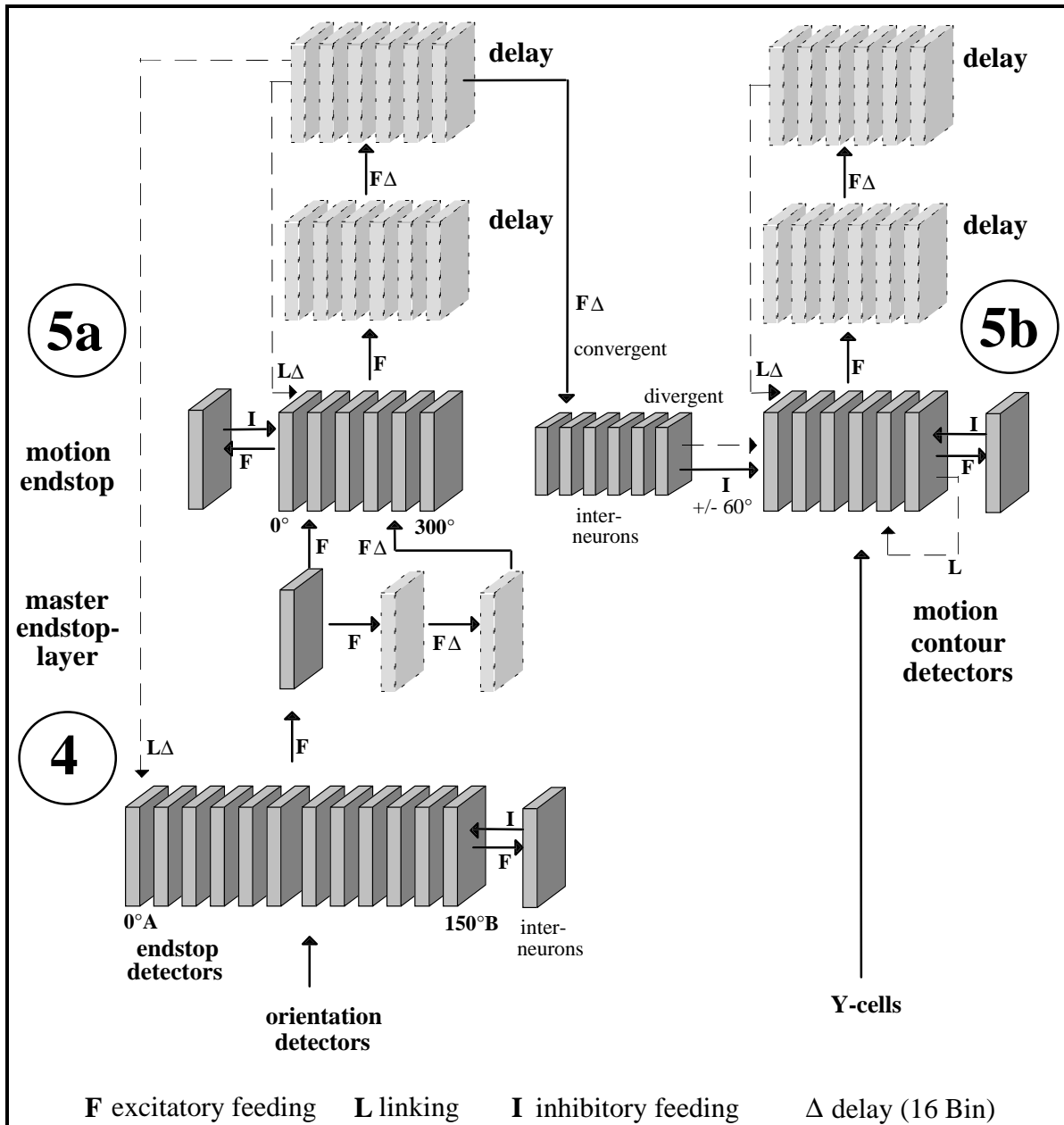


Abb. 3.84: Modell zur Einwirkung der Linienendenbewegung auf die Konturbewegung. In 4 werden die Linienenden extrahiert und nachfolgend in einer Schicht zusammengefasst. Zur Bewegungsrichtungsbestimmung werden die die Linienenden repräsentierenden neuronalen Aktivitäten um 32 Bin verzögert und vorwärts in Bewegungsrichtung übertragen. Die extrahierte Bewegungsrichtung der Linienenden (5a) wird dann über Interneuronen an die Richtungsdetektoren für Konturen (5b) gegeben. Die Verwendung von Interneuronen anstelle direkter Verbindungen reduziert die Anzahl an Verbindungen erheblich.

Die Richtungsdetektoren für Linienenden sind wiederum miteinander gekoppelt. Allerdings macht hier nur eine verzögerte Kopplung zur Voraktivierung des wahrscheinlich nächsten vom Linienende stimulierten Detektors Sinn. Eine laterale, unverzögerte Kopplung ist unnötig, da in der unmittelbaren Umgebung eines Linienendes im Eingabebild kein weiteres Linienende zu erwarten ist. Um eine eindeutige Bewegungsrichtung und eine Aktivitätsbeschränkung zu gewährleisten, sind die Richtungsdetektoren für Linienenden wiederum über inhibitorisch wirkende Interneuronen gekoppelt. Die Verschaltung der Voraktivierung und der Interneuronen erfolgt in Analogie zu der in Kapitel 3.1 beschriebenen.

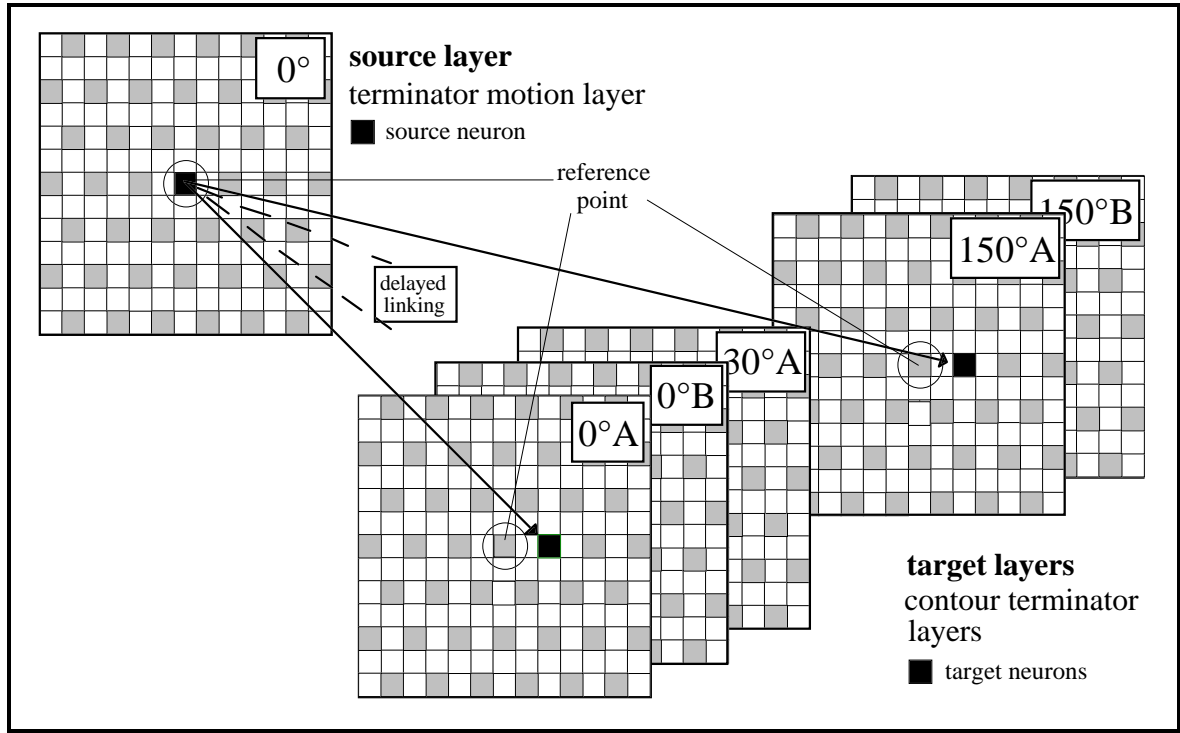


Abb. 3.85: Voraktivierung der Linienendendetektoren durch Richtungsdetektoren für Linienenden. Durch die Bestimmung der Bewegungsrichtung der Linienenden kann deren nächste Position im Bild geschätzt werden. Durch eine Voraktivierung der entsprechenden Linienendendetektoren wird das Detektionsergebnis verbessert. Die Abbildung zeigt die Voraktivierung am Beispiel eines 0°-LEBRD der feinen Auflösung ($d' = 2$ Pixel).

Die von den LEBRD bestimmte Bewegungsrichtung kann zusätzlich zur Voraktivierung der Linienendendetektoren in Bewegungsrichtung benutzt werden. In Abbildung 3.85 ist als Beispiel die Voraktivierung der Linienendendetektoren durch die 0°-LEBRD dargestellt. Da die Orientierungsinformation in der Linienenden-Master-Schicht verlorengeht, kann nicht mehr orientierungsspezifisch voraktiviert werden. Vielmehr müssen alle Linienendendetektoren, deren Abtastort mit der geschätzten Position des Linienendes im nächsten Bild übereinstimmt, über Linking-Verbindungen voraktiviert werden. Die Verschaltung dieser Voraktivierung wird durch Gleichung 3.82 beschrieben.

$$M\text{-}ESDDb,0^\circ_{\text{link}} w_{ijkl}^{ESD,\phi A/B} = \begin{cases} {}^{ESD}V^L, & \text{falls } ((i == k) \wedge (j == (l + d'))) \\ 0, & \text{sonst} \end{cases} \quad (\text{Gl. 3.82})$$

mit $\phi \in \{0^\circ, 30^\circ, 60^\circ, 90^\circ, 120^\circ, 150^\circ\}$

Nachdem die Bewegungsrichtungen der Linienenden bestimmt worden sind, sollen sie auf die Richtungsdetektoren der Kontur Einfluß nehmen. Dazu muß die Information der Linienendenbewegungsrichtung langreichweitig auf die Richtungsdetektoren übertragen werden. Damit die durch die Richtungsdetektoren repräsentierten Konturen erhalten bleiben, darf die Beeinflussung nicht durch Feeding-Verbindungen erfolgen, denn diese würden alle benachbarten Richtungsdetektoren aktivieren und die Konturinformation ginge verloren.

Auch hier bietet sich eine Kopplung über Linking-Verbindungen an. Gleichzeitig müssen die Richtungsdetektoren, deren Richtungspräferenz ungleich der extrahierten Linienendenbewegungsrichtung ist, gehemmt werden. Eine direkte Kopplung der LEBRD mit den KBRD würde zu einer zu großen Anzahl an Verbindungen führen. Deswegen werden zusätzliche

Interneuronen eingeführt, die die Information über die Linienendenbewegungsrichtung aufnehmen und an die KBRD weitergeben. Die Abbildung 3.86 zeigt die konvergente Verschaltung der LEBRD, deren Ausgabe um 32 Bin verzögert wurde, zu den Interneuronen. Zur Informationsübermittlung sind nicht viele Interneuronen notwendig, so daß die Abtastweite der Interneuronen viermal so groß gewählt werden kann wie die der Richtungsdetektoren für Linienenden.

Für alle sechs Schichten mit Richtungsdetektoren für Linienenden existiert jeweils eine Schicht mit Interneuronen, so daß die Verschaltung mit Gleichung 3.83 beschrieben werden kann.

$${}_{\text{feed1}}^{M\text{-ESDDb}, \phi' = \phi} w_{ijkl}^{M\text{-ESD}, \text{Inter } \phi} = \begin{cases} {}^{\text{IM-ESD}} V_n^{F1}, & \text{für hexagon}_d n \\ 0, & \text{sonst} \end{cases} \quad (\text{Gl. 3.83})$$

mit $\phi \in \{0^\circ, 60^\circ, 120^\circ, 180^\circ, 240^\circ, 300^\circ\}$ und $0 \leq n \leq 7$

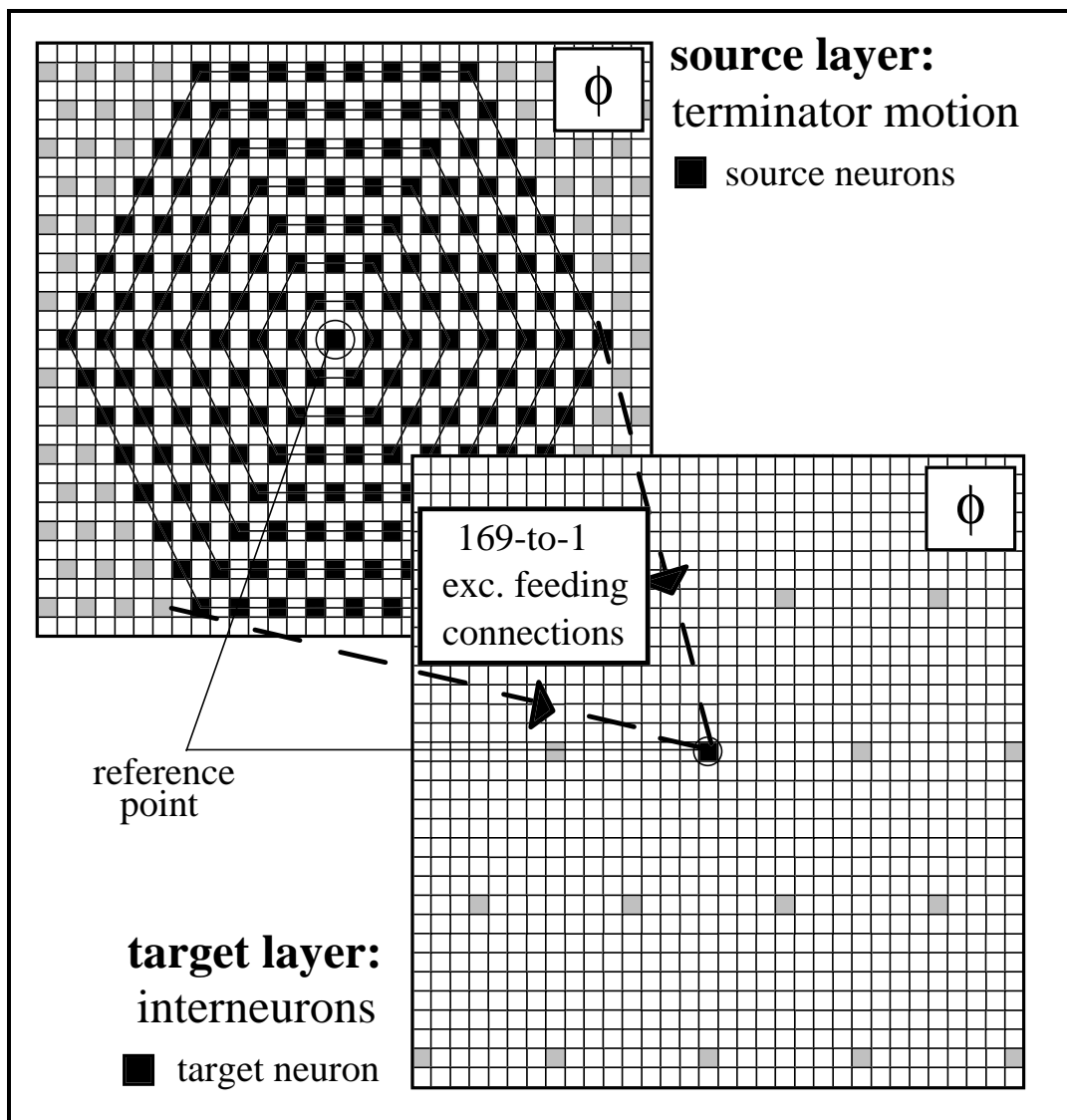


Abb. 3.86. Konvergente Verschaltung der Richtungsdetektoren für Linienenden auf Interneuronen. Ein Interneuron sammelt die Ausgabe von 169 LEBRD ein. Dargestellt sind hierbei Detektoren der Abtastweite $d = 2$ Pixel. Bei der mittleren Auflösung verdoppelt sich der Abstand der Interneuronen entsprechend.

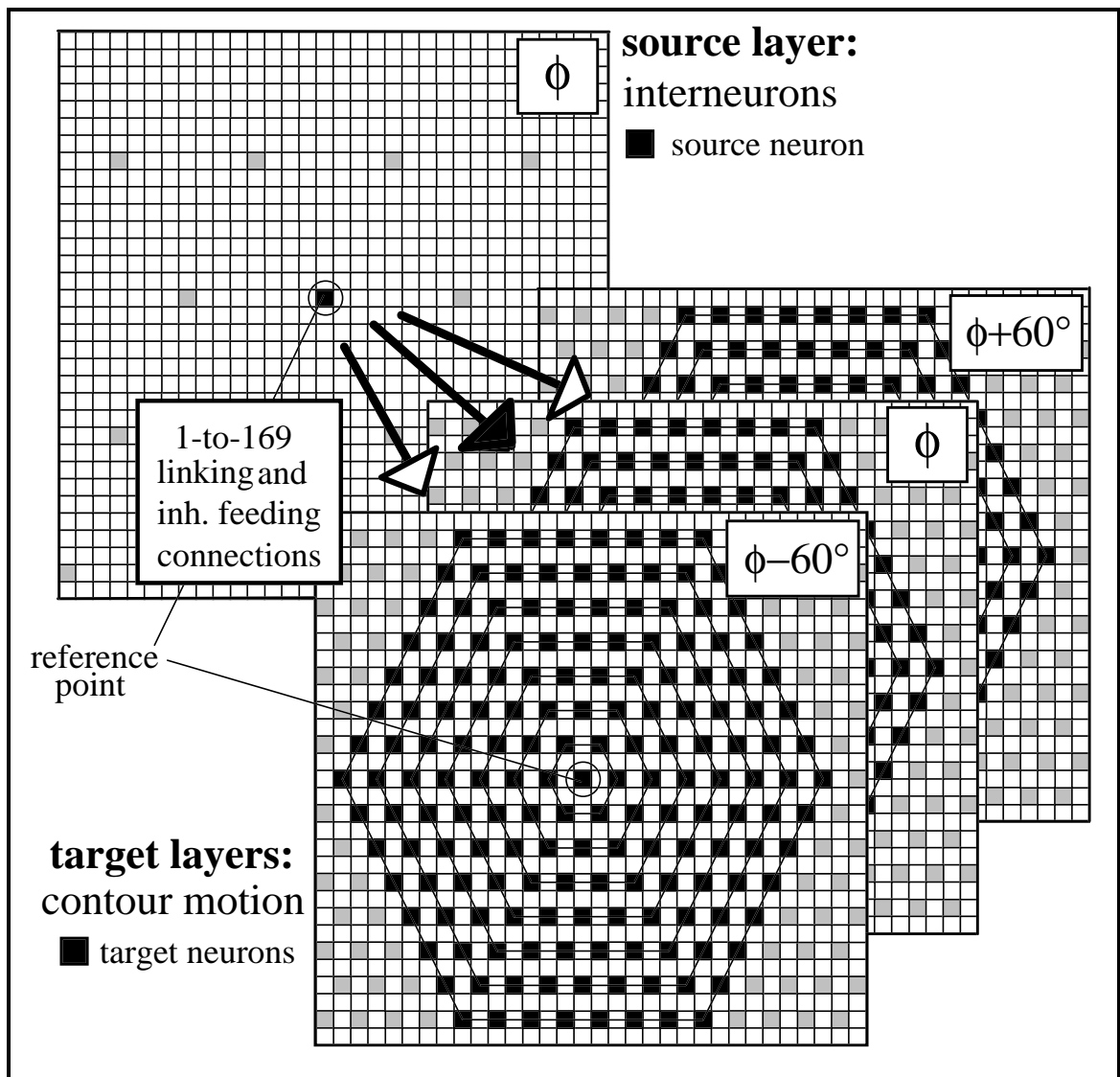


Abb. 3.87: Divergente Verschaltung der Interneuronen auf Richtungsdetektoren für Konturen. Ein Interneuron, das seine Eingabe über LEBRD mit einer Richtungspräferenz von ϕ erhält, gibt seine Ausgabe divergent an 169 KBRD der gleichen Richtungspräferenz über Linking-Verbindungen weiter. Weiterhin werden von diesem Interneuron jeweils 169 Modellneuronen, deren Richtungspräferenz um $\pm 60^\circ$ von ϕ verschieden ist, gehemmt.

Die Verbindungsstärke ${}^{\text{IM-ESD}}V_n^{\text{F1}}$ fällt hierbei linear mit dem Abstand zwischen den beiden zu koppelnden Modellneuronen ab. Die genauen Werte sind der Tabelle 3.8 zu entnehmen.

Die Interneuronen übertragen, wie in Abbildung 3.87 dargestellt, die Bewegungsrichtung der Linienenden auf die KBRD. Da in den Schichten der LEBRD keine Information mehr über die Orientierung der Linienenden enthalten ist, erfolgt die Kopplung isotrop. Selbstverständlich könnte auch die Orientierung anhand der Aktivität in den Linienendendetektorschichten ermittelt werden, um eine gezielte, anisotrope Kopplung möglich zu machen. Dafür wären aber wiederum eine große Anzahl an Modellneuronen nötig, die als Kontroll- oder Schaltereinheiten fungieren müssten. Aus diesem Grund findet eine isotrope Kopplung der Interneuronen mit den Richtungsdetektoren gemäß den Gleichungen 3.84 und 3.85 statt. Die Verbindungsstärken ${}^{\text{M}}V_n^{\text{L}}$ und ${}^{\text{M}}V_n^{\text{I}}$ fallen abstandsabhängig ab.

$$w_{ijkl}^{M-ESD, \phi} = \begin{cases} V_n^L, & \text{für hexagon}_d n \\ 0, & \text{sonst} \end{cases} \quad 0 \leq n \leq 7 \quad (\text{Gl. 3.84})$$

$$w_{ijkl}^{M-ESD, \phi} = \begin{cases} V_n^{F1}, & \text{für hexagon}_d n \\ 0, & \text{sonst} \end{cases} \quad 0 \leq n \leq 7 \quad (\text{Gl. 3.85})$$

mit $\phi \in \{0^\circ, 60^\circ, 120^\circ, 180^\circ, 240^\circ, 300^\circ\}$

Orientierungsdetektoren	
Feeding-Gewicht Orientierungsdetektor $^{OD}V^{F2}$	1,0
Feeding-Zeitkonstante $^{OD}\tau^{F2}$	5,0 Bin
Linking-Gewicht $^{OD}V^L$	1,0
Linking-Zeitkonstante $^{OD}\tau^L$	5,0 Bin
Inhibition-Gewicht $^{OD}V^I$	5,0
Inhibition-Zeitkonstante $^{OD}\tau^I$	10,0 Bin
Schwellenverstärkung Orientierungsdetektor $^{OD}V^\Theta$	16
Schwellenzeitkonstante $^{OD}\tau^\Theta$	3,0 Bin
Schwellenoffset $^{OD}\Theta_0$	4
Feeding-Gewicht Inhibitionsneuron $^{IOD}V^{F1}$	3,0
Feeding-Zeitkonstante $^{IOD}\tau^{F1}$	5,0
Schwellenverstärkung Inhibitionsneuron $^{IOD}V^\Theta$	16
Schwellenzeitkonstante $^{IOD}\tau^\Theta$	5,0
Schwellenoffset $^{IOD}\Theta_0$	8
Orientierungsdetektoren II	
Feeding-Gewicht Orientierungsdetektor II $^{OD-II}V^{F2}$	3,5
Feeding-Zeitkonstante $^{OD-II}\tau^{F2}$	5 Bin
Linking-Gewicht $^{OD-II}V^L$	1,0
Linking-Zeitkonstante $^{OD-II}\tau^L$	5 Bin
Inhibition-Gewicht $^{OD-II}V^I$	2,5
Inhibition-Zeitkonstante $^{OD-II}\tau^I$	10 Bin
Schwellenverstärkung Orientierungsdetektor II $^{OD-II}V^\Theta$	0
Schwellenzeitkonstante $^{OD-II}\tau^\Theta$	3 Bin
Schwellenoffset $^{OD-II}\Theta_0$	4
Feeding-Gewicht Inhibitionsneuron $^{IOD-II}V^{F1}$	3,0
Feeding-Zeitkonstante $^{IOD-II}\tau^{F1}$	5 Bin
Schwellenverstärkung Inhibitionsneuron $^{IOD-II}V^\Theta$	32
Schwellenzeitkonstante $^{IOD-II}\tau^\Theta$	5 Bin
Schwellenoffset $^{IOD-II}\Theta_0$	8
Linienendendetektoren	
Feeding-Gewicht Linienendendetektor $^{ESD}V^{F2}$	0,5
Feeding-Zeitkonstante $^{ESD}\tau^{F2}$	5 Bin
Linking-Gewicht $^{ESD}V^L$	3,0
Linking-Zeitkonstante $^{ESD}\tau^L$	5 Bin

Inhibition-Gewicht ${}^{\text{ESD}}V^I, {}^{\text{ESD}}V^{FI}$	7,0, 7,5
Inhibition-Zeitkonstante ${}^{\text{ESD}}\tau^I$	10 Bin
Schwellenverstärkung Linienendendetektor ${}^{\text{ESD}}V^\Theta$	8
Schwellenzeitkonstante ${}^{\text{ESD}}\tau^\Theta$	3 Bin
Schwellenoffset ${}^{\text{ESD}}\Theta_0$	8
Feeding-Gewicht Inhibitionsneuron ${}^{\text{IESD}}V^{FI}$	7,5
Feeding-Zeitkonstante ${}^{\text{IESD}}\tau^{FI}$	5 Bin
Schwellenverstärkung Inhibitionsneuron ${}^{\text{IESD}}V^\Theta$	16
Schwellenzeitkonstante ${}^{\text{IESD}}\tau^\Theta$	5 Bin
Schwellenoffset ${}^{\text{IESD}}\Theta_0$	8
Master-Linienendendetektoren	
Feeding-Gewicht Master-Linienendendetektor ${}^{\text{ESD,Master}}V^{FI}$	4,5
Feeding-Zeitkonstante ${}^{\text{ESD,Master}}\tau^{FI}$	5 Bin
Schwellenverstärkung Master-Linienendendetektor ${}^{\text{ESD,Master}}V^\Theta$	8
Schwellenzeitkonstante ${}^{\text{ESD,Master}}\tau^\Theta$	3 Bin
Schwellenoffset ${}^{\text{ESD,Master}}\Theta_0$	4
Richtungsdetektoren für Linienenden (LEBRD)	
Feeding-Gewicht LEBRD ${}^{\text{M-ESD}}V^{FI}$	2,5
Feeding-Zeitkonstante ${}^{\text{M-ESD}}\tau^{FI}$	5 Bin
Linking-Gewicht ${}^{\text{M-ESD}}V^L$	5,0
Linking-Zeitkonstante ${}^{\text{M-ESD}}\tau^L$	5 Bin
Inhibition-Gewicht ${}^{\text{M-ESD}}V^I$	5,0
Inhibition-Zeitkonstante ${}^{\text{M-ESD}}\tau^I$	10 Bin
Schwellenverst. LEBRD ${}^{\text{M-ESD}}V^\Theta$	8
Schwellenzeitkonstante ${}^{\text{M-ESD}}\tau^\Theta$	3 Bin
Schwellenoffset ${}^{\text{M-ESD}}\Theta_0$	4
Feeding-Gewicht Inhibitionsneuron ${}^{\text{IM-ESD}}V^{FI}$	7,5
Feeding-Zeitkonstante ${}^{\text{IM-ESD}}\tau^{FI}$	5 Bin
Schwellenverstärkung Inhibitionsneuron ${}^{\text{IM-ESD}}V^\Theta$	8
Schwellenzeitkonstante ${}^{\text{IM-ESD}}\tau^\Theta$	5 Bin
Schwellenoffset ${}^{\text{IM-ESD}}\Theta_0$	8
Feeding-Gewicht Interneuron ${}^{\text{M-ESD,Inter}}V_n^{FI}$ (mit $0 \leq n \leq 7$)	$4-n/2$
Feeding-Zeitkonstante ${}^{\text{M-ESD,Inter}}\tau^{FI}$	5 Bin
Schwellenverstärkung Interneuron ${}^{\text{M-ESD,Inter}}V^\Theta$	32
Schwellenzeitkonstante ${}^{\text{M-ESD,Inter}}\tau^\Theta$	5 Bin
Schwellenoffset ${}^{\text{M-ESD,Inter}}\Theta_0$	8
Richtungsdetektoren (zusätzliche Verbindungen)	
Linking-Gewicht Richtungsdetektor ${}^{\text{M}}V_n^L$ (mit $0 \leq n \leq 7$)	7,5-n
Feeding-Gewicht Richtungsdetektor ${}^{\text{M}}V_0^{FI}, {}^{\text{M}}V_1^{FI}, {}^{\text{M}}V_2^{FI}$	-7,5, -7,5, -7,25
Feeding-Gewicht Richtungsdetektor ${}^{\text{M}}V_3^{FI}, {}^{\text{M}}V_4^{FI}, {}^{\text{M}}V_5^{FI}$	-7,0, -6,0, -4,5
Feeding-Gewicht Richtungsdetektor ${}^{\text{M}}V_6^{FI}, {}^{\text{M}}V_7^{FI}$	-2,5, -1,25

Tab. 3.8: Parameter der Orientierungs- und Linienendendetektoren sowie Richtungsdetektoren für Konturen und Linienenden.

3.4.2 Simulationen zur Einbindung der Linienenden in die Bewegungsverarbeitung

3.4.2.1 Simulation der Barberpole Illusion und der Barberplaid Illusion

In diesem Kapitel werden sowohl die *barberpole illusion* als auch die *barberplaid illusion* mit dem in Kapitel 3.4.1 beschriebenen neuronalen Netzwerk simuliert. Zunächst wird jedoch das neuronale Netzwerk aus den Abbildungen 3.78 und 3.84 stufenweise erläutert. Dazu werden die Eigenschaften der Orientierungsdetektoren dargestellt und die Funktionalität der Linienendendetektoren sowie der Richtungsdetektoren für Linienenden (LEBRD) an einer künstlichen Bildsequenz erklärt. Im nachfolgenden Kapitel 3.4.2.2 wird dann auch die Einsetzbarkeit des neuronalen Netzwerks bei Eingabe realer Szenen gezeigt.

Die Verschaltung der Orientierungsdetektoren wurde so gewählt, daß anhand der neuronalen Aktivität der Modellneuronen eine eindeutige Bestimmung der Kantenorientierung ohne Beeinflussung durch die Kantenbewegungsrichtung vorgenommen werden kann. Dazu muß das rezeptive Feld der Orientierungsdetektoren symmetrisch aufgebaut sein. Zur Untersuchung der Orientierungscharakteristiken der Orientierungsdetektoren wurde ein Netz mit 20 Abtastzeilen und 20 Abtastspalten sowie einer Abtastweite von $d = 4$ Pixel benutzt. Als Stimulus diente eine verrauschte ($V^{\text{noise}} = 10$) Intensitätskontrastkante mit den Grauwerten $g_1 = 150$ und $g_2 = 50$, die über das Abtastraster mit sechs unterschiedlichen Geschwindigkeiten ($1 \text{ Pixel/Bild} \leq v \leq 6 \text{ Pixel/Bild}$) orthogonal zur Kantenorientierung bewegt wurde.

In den Abbildungen 3.88 und 3.89 sind die über zehn Durchläufe mit den jeweils für die Geschwindigkeit spezifischen Positionsverschiebungen gemittelten Aktivitäten der ON- bzw. OFF-Orientierungsdetektoren eines Orientierungsdetektorsatzes gegen die Bewegungsrichtung der Intensitätskontrastkante aufgetragen. Da es sich bei diesem Stimulus um eine sich bewegende hell/dunkel-Kante handelt, generieren die Modellneuronen des ON-Pfades eine stärkere Aktivität als die des OFF-Pfades. Die Verarbeitung im ON-Pfad zeigt, daß die Modellneuronen bei einer Kantengeschwindigkeit von $v = 3 \text{ Pixel/Bild}$ am aktivsten sind. Bei einer Kantengeschwindigkeit von $v = 1 \text{ Pixel/Bild}$ werden die Y-Zellen der Vorverarbeitung noch nicht ausreichend stimuliert, so daß auch die Orientierungsdetektoren nur eine schwache Aktivierung aufweisen. Ab einer Kantengeschwindigkeit von $v = 2 \text{ Pixel/Bild}$ werden alle Kanten von den entsprechenden Modellneuronen, deren Orientierungspräferenz der Kantenorientierung ähnelt, detektiert, wobei unter einer Geschwindigkeit von $v = 4 \text{ Pixel/Bild}$ die Orientierungsdetektion noch mehrdeutig ist. Für größere Geschwindigkeiten ist die Detektion eindeutig, es sei denn, die Kantenorientierung liegt genau zwischen zwei Orientierungspräferenzen. Da in diesem Fall beide Modellneuronen gleich stark stimuliert werden, kann keines den lokalen Wettstreit gewinnen, wodurch beide aktiv sind.

Die Orientierungsdetektoren des OFF-Pfades sind erst ab einer Kantengeschwindigkeit von $v = 3 \text{ Pixel/Bild}$ aktiv. Die Orientierungsdetektion ist wie die des ON-Pfades eindeutig, aber auf einem deutlich kleineren Aktivitätsniveau.

Da die den Orientierungsdetektoren nachgeschalteten ONOFF-Orientierungsdetektoren keine funktionelle Bedeutung haben - sie dienen lediglich zur Reduzierung der Anzahl an notwendigen Linienendendetektoren -, wird hier auf eine genauere Untersuchung verzichtet.

Die Extraktion der Linienenden und deren Bewegungsrichtung wird zunächst anhand eines Beispiels mit einem künstlich generierten Stimulus dargestellt. Im nachfolgenden Kapitel 3.4.2.2 wird dann die Extraktion der Linienenden und deren Bewegungsrichtung anhand realer Szenen demonstriert.

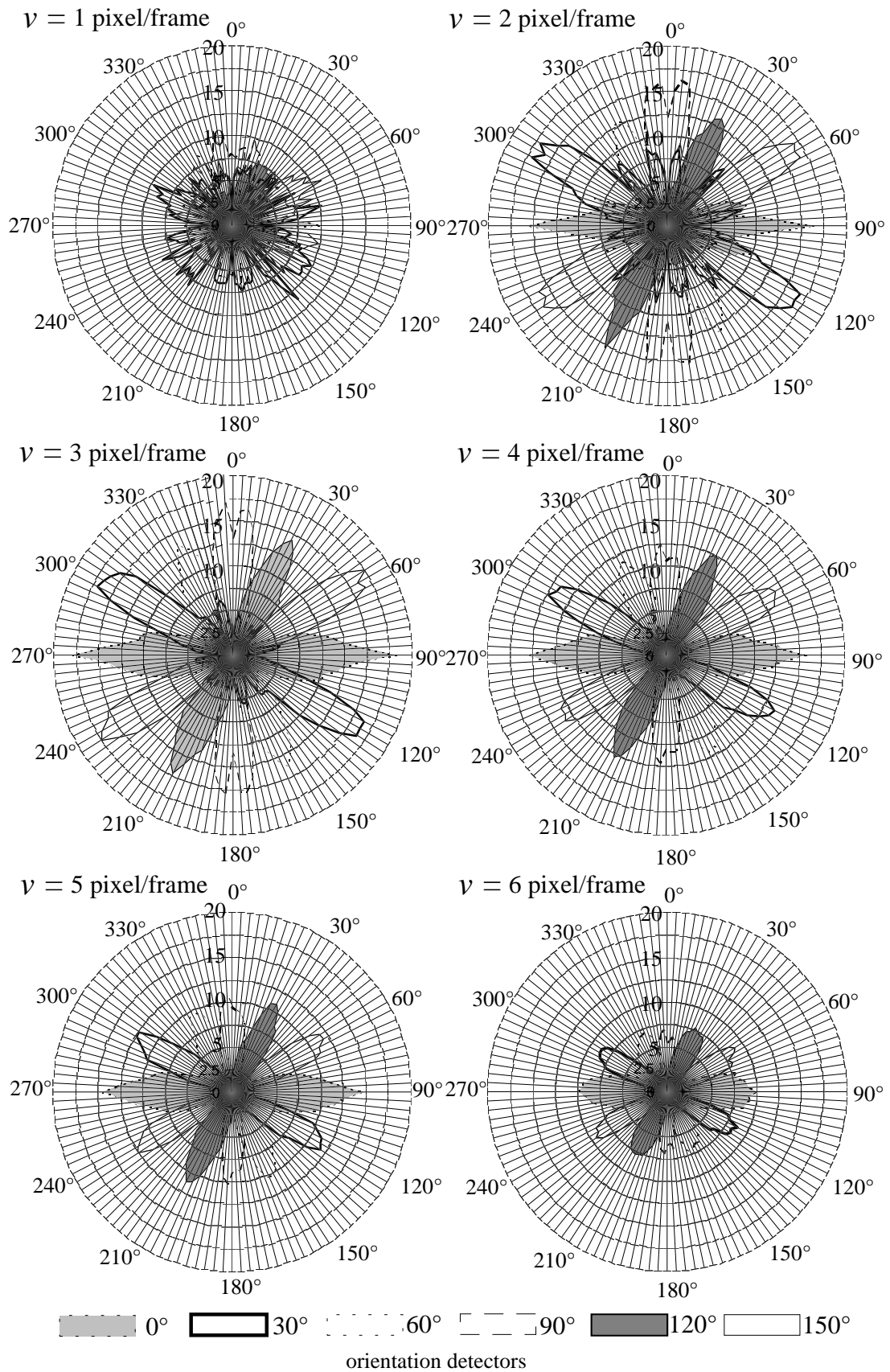


Abb. 3.88: Orientierungscharakteristik der ON-Orientierungsdetektoren. Die Abbildung stellt die von den ON-Orientierungsdetektoren der mittleren Auflösung generierte Anzahl an Aktionspotentialen pro Kantendurchlauf in Abhängigkeit von der Bewegungsrichtung der Intensitätskontrastkante orthogonal zur Kantenorientierung für sechs verschiedene Geschwindigkeiten dar.

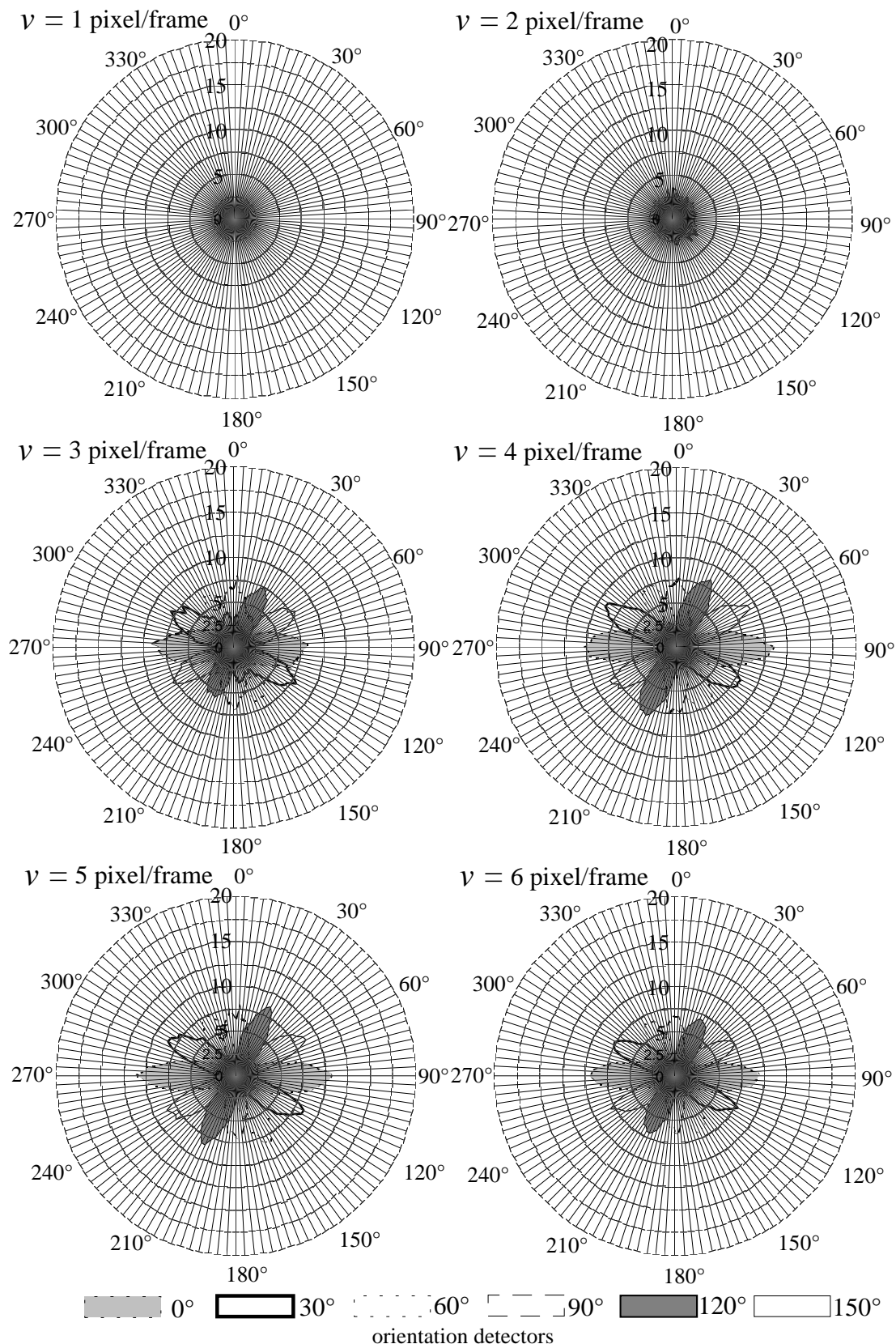


Abb. 3.89: Orientierungscharakteristik der OFF-Orientierungsdetektoren. Die Abbildung stellt die von den OFF-Orientierungsdetektoren der mittleren Auflösung generierte Anzahl an Aktions-potentialen pro Durchlauf in Abhängigkeit von der Bewegungsrichtung der Intensitätskontrastkante orthogonal zur Kantenorientierung für sechs verschiedene Geschwindigkeiten dar. Die Aktivitäten der OFF-Orientierungsdetektoren sind geringer als die der ON-Orientierungsdetektoren, da mit einer hell/dunkel-Kante stimuliert wurde (vgl. Kapitel 3.1).

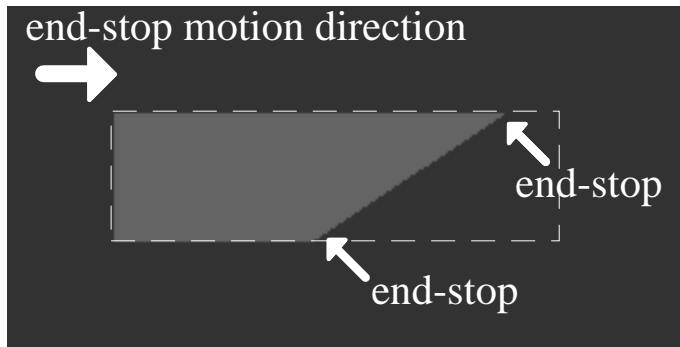


Abb. 3.90: Stimulus zur Untersuchung der Linienendenextraktion und der Linienendenbewegungsrichtungsdetektion. Dargestellt ist eine 30° orientierte Kante, die hinter einer Verdeckung mit einer rechteckigen Apertur bewegt wird. Dadurch entstehen zwei Linienenden (Terminatoren), die sich in Richtung 0° bewegen. Die gestrichelte Linie dient lediglich als Markierung der Apertur und ist im Stimulus nicht enthalten.

Als Eingabe wurde der in der Abbildung 3.90 dargestellte Stimulus verwendet. Es handelt sich hierbei um eine 30° orientierte Intensitätskontrastkante ($g_1 = 150$, $g_2 = 50$), die hinter einer Verdeckung ($g = 50$) mit einer rechteckigen Apertur in Richtung 300° bewegt wird. Durch die partielle Verdeckung entstehen Linienenden, die sich in Richtung 0° bewegen. Das Netz bestand aus 20 Abtastzeilen und 52 Abtastspalten, und es wurde eine Abtastweite von $d = 4$ Pixel verwendet. Die Linienendenbewegungsrichtungsdetektoren erhalten in dieser Simulation ihre Eingabe von zwei Modellneuronen der Linienenden-Masterschicht, deren RF-Mittelpunkte $d' = 8$ Pixel voneinander entfernt sind.

Das Ergebnis der Linienendenextraktion ist in Abbildung 3.91 für fünf aufeinanderfolgende Bilder der Eingabesequenz dargestellt. Die linke Spalte dieser Abbildung zeigt die Aktivität der 30° -Linienendendetektoren A und die rechte Spalte die der 30° -Linienendendetektoren B. Die Aktivität der Linienendendetektoren wurde hierbei als rezeptives Feld der retinalen Ganglienzellen desselben Abtastortes in das Eingabebild zurückprojiziert (Kopecz et al. 1996). Die Abbildung 3.91 zeigt die fehlerfreie Bestimmung der Linienenden durch die 30° -Linienendendetektoren. Die Linienenden auf der linken Seite des Stimulus werden nicht detektiert, da sie nicht bewegt und damit bereits von den Y-Zellen nicht detektiert werden.

Da die Bestimmung der Linienendenbewegungsrichtung ein aufwendiger Prozeß ist, wird sie an dem unteren Linienende des Stimulus beispielhaft erläutert. Die Verschaltung ausgewählter Modellneuronen und deren Aktivität ist in Abbildung 3.92 aufgetragen. Die Ausgaben der Linienendendetektoren werden von der Linienenden-Master-Schicht gesammelt. Dies dient lediglich der Reduzierung der Modellneuronenanzahl. Ohne diese zusätzliche Master-Schicht müßte zu jeder Linienendendetektorschicht ein kompletter Satz Richtungsdetektoren zur Verfügung gestellt werden. Dadurch würde sich zwar das Korrespondenzproblem verringern, aber die Modellneuronenanzahl würde sich um den Faktor 12 vervielfachen. In der Abbildung 3.92 wird bereits die Geschwindigkeit der Linienenden an der Aktivität der Linienendendetektoren sichtbar. Die Linienendengeschwindigkeit von 8 Pixel/Bild führt zu einer starken Aktivierung jedes zweiten Linienendendetektors (Modellneuronen 1, 3 und 5). Die Modellneuronen 2 und 4 sind schwach aktiv, da das Linienende diese geringer aktiviert, was auf eine größere Entfernung des Linienendes zu den korrespondierenden Abtastorten der Modellneuronen zurückzuführen ist. Dieser Aktivitätsunterschied ist auch noch in der Aktivität der Master-Schicht-Modellneuronen 6, 7 und 8 zu erkennen. Um die Bewegungsrichtung der Linienenden bestimmen zu können, muß eine Detektorantwort verzögert und die Antwort eines benachbarten Detektors unverzögert auf den LEBRD gegeben werden. In der Abbildung 3.92 wird dies an den Modellneuronen 6, 8, 9, 10 und 11 deutlich. Modellneuron 11 ist ein LEBRD. Damit dieses Modellneuron sensitiv für eine Bewegungsrichtung von 0° wird, muß es eine unverzögerte Eingabe von Modellneuron 8 erhalten. Da sich die Linienenden mit einer Geschwindigkeit von 8 Pixel/Bild bewegen, muß der zweite Eingang des LEBRD vom Abtastort des sechsten Modellneurons kommen. Die Verzögerung der Ausgabe des sechsten

Modellneurons wird über die Modellneuronen 9 und 10 realisiert. Zur Verbesserung der Detektionseigenschaften der LEBRD wird in Analogie zu den KBRD eine Voraktivierung implementiert. Dazu muß die Ausgabe des elften Modellneurons über die Modellneuronen 12 und 13 verzögert werden. Die Ausgabe des 13. Modellneurons kann dann sowohl den 0°-LEBRD als auch die Linienendendetektoren an dem geschätzten Abtastort über Linking-Verbindungen voraktivieren. Die primäre Aufgabe der verzögerten Aktivität der LEBRD ist jedoch die Beeinflussung der KBRD, wie es im folgenden in den Simulationen der *barberpole illusion* und der *barberplaid illusion* gezeigt wird.

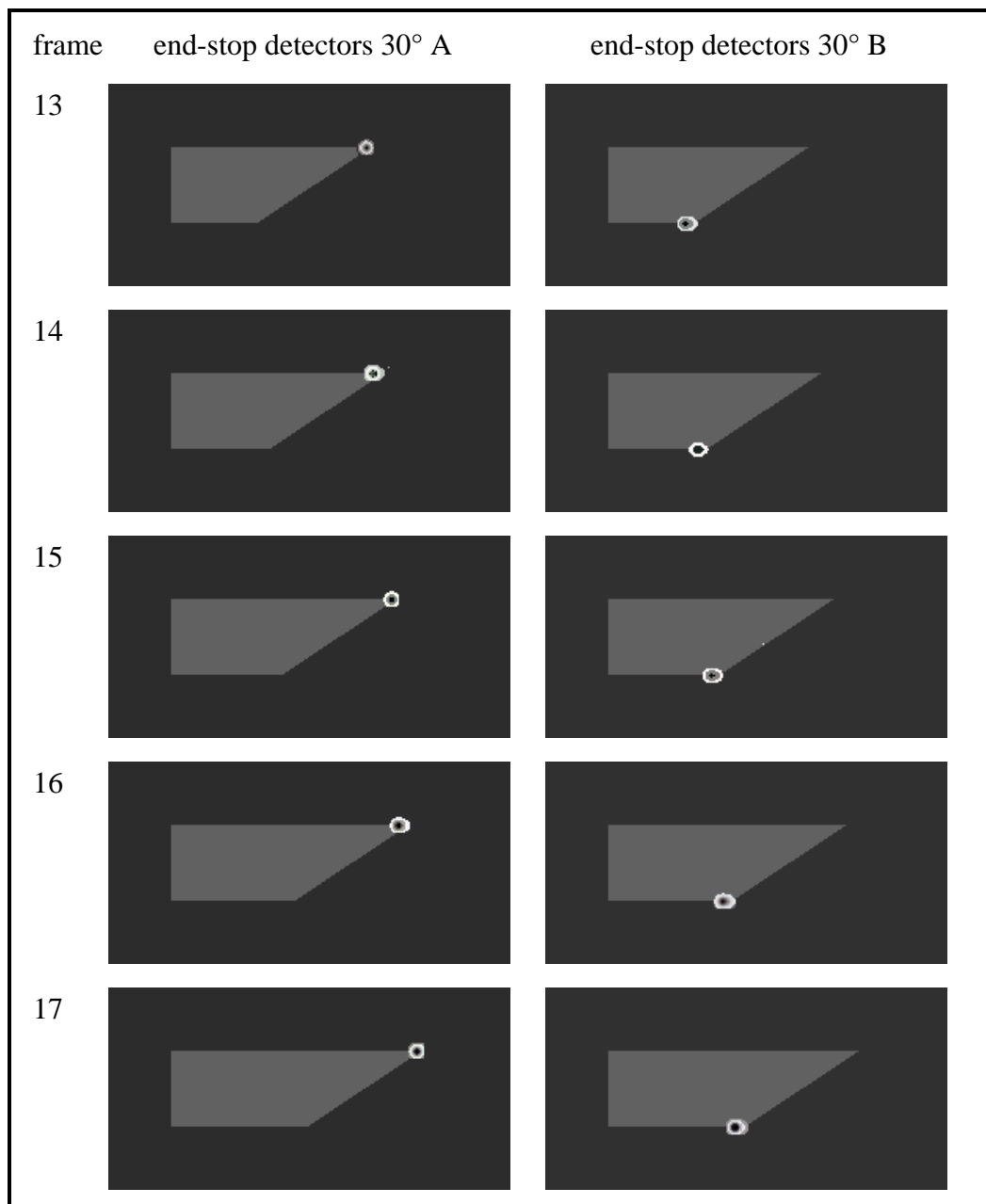


Abb. 3.91: Simulationsergebnis Linienendenextraktion. Die Abbildung zeigt die Aktivität der 30°-Linienendendetektoren in fünf aufeinanderfolgenden Bildern der Eingabesequenz. Die Linienenden bewegen sich in diesem Beispiel mit $v = 8$ Pixel/Bild in Richtung 0°.

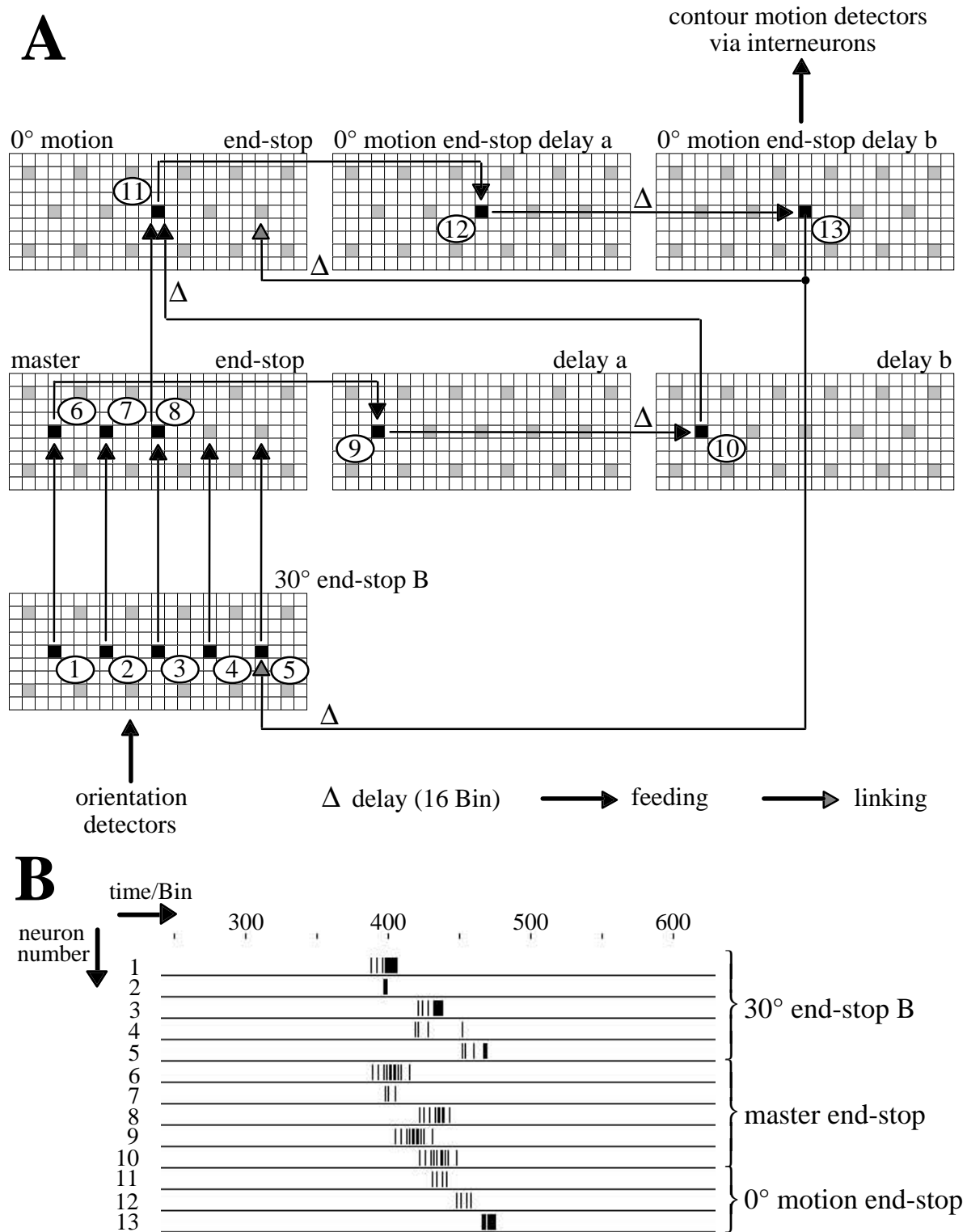


Abb. 3.92: Extraktion der Linienendenbewegungsrichtung. Verschaltungsausschnitt (A) zur Bestimmung der Bewegungsrichtung der Linienenden anhand des Simulationsbeispiels (B) eines 30° orientierten Linienendes, das sich in Richtung 0° bewegt. A: Die Linienendendetektoren erhalten ihre Eingabe von den Orientierungsdetektoren. Eine Master-Schicht sammelt zur Verringerung der Modellneuronenanzahl in den nachfolgenden Verarbeitungsstufen über eine konvergente Verschaltung die neuronale Aktivität der Linienendendetektoren ein. Damit eine Bewegungsrichtungsdetektion möglich wird, müssen die Ausgaben der Linienenden-Master-Schicht zum einen unmittelbar, zum anderen um 32 Bin verzögert auf die LEBRD gegeben werden. Das Ergebnis der Richtungsdetektion der Linienenden wird einerseits über Interneurone an die KBRD weitergeleitet, andererseits dient es zur Voraktivierung der Linienendendetektoren und der LEBRD. Die anderen LEBRD-Schichten, die Schichten der übrigen Linienendendetektoren sowie deren Verschaltung untereinander und über Interneurone sind aus Gründen der Übersichtlichkeit nicht dargestellt.

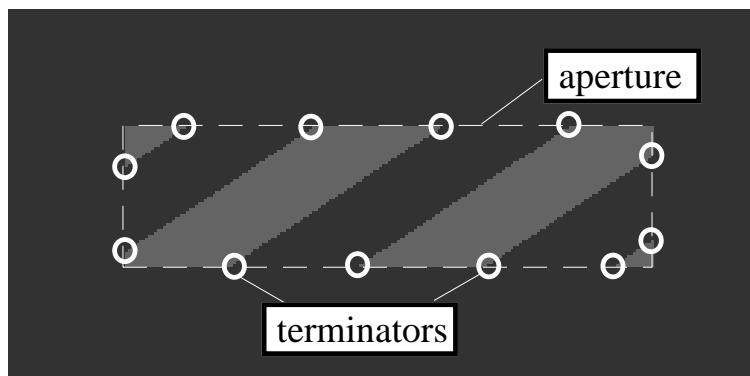
Barberpole Illusion

Zur Simulation der *barberpole illusion* wurde der in Abbildung 3.93 dargestellte Stimulus benutzt. Die 30° orientierten Balken des Gitters hatten Grauwerte von $g_1 = 100$ und $g_2 = 50$ und wurden hinter einer Verdeckung ($g = 50$) mit einer rechteckigen Apertur mit unterschiedlichen Geschwindigkeiten in Richtung 300° bewegt. Das neuronale Netzwerk bestand wiederum aus 52 Abtastspalten und 20 Abtastzeilen (Abtastweite $d = 4$ Pixel).

Die Abbildung 3.94 zeigt die Aktivität der ON- und OFF-Richtungsdetektoren der mittleren Auflösung ohne (A) und mit (B) Interaktion der LEBRD mit den KBRD bei Eingabe des 40. Bildes der Sequenz. Die Terminatoren bewegen sich in dieser Sequenz mit einer Geschwindigkeitskomponente in Richtung 0° von $v_{0^\circ} = 8$ Pixel/Bild (Positionsverschiebung $\Delta = 1$ Pixel). Ohne Interaktion sind hauptsächlich die 300° -KBRD aktiv. Die KBRD signalisieren damit eine Bewegung orthogonal zur Balkenorientierung. Diese Bewegungsrichtung nehmen menschliche Beobachter wahr, wenn sie eine sich bewegende Kante durch eine runde Apertur beobachten, wobei zu bemerken ist, daß es bei einer runden Apertur keine gemeinsame Linienendenbewegungsrichtung gibt. Im oberen Teil des Eingabebildes wird auch vereinzelt eine Bewegungsrichtung in 0° detektiert. Die 0° -Richtungsdetektoren finden aber im Gegensatz zu den 300° -Richtungsdetektoren kaum Unterstützung durch benachbarte Detektoren gleicher Richtungspräferenz, wodurch ihre Aktivität gering bleibt. Die 300° -KBRD, deren rezeptive Felder sich entlang der 30° orientierten Kontur befinden, gewinnen den Wettbewerb gegen die 0° - und 240° -KBRD, da sie sich wechselseitig über Linking-Verbindungen unterstützen können (vgl. Kapitel 3.2.1, Abbildung 3.31). Damit die 0° -KBRD diesen lokalen Wettbewerb gewinnen können, benötigen sie Unterstützung von den 0° -LEBRD. Diese erhalten sie über die weit ausgreifenden Linking-Verbindungen von den LEBRD zu den KBRD. Zusätzlich werden die Richtungsdetektoren, deren Richtungspräferenz um $\pm 60^\circ$ von der Richtungspräferenz der LEBRD abweicht, gehemmt. In diesem konkreten Fall inhibieren somit die 0° -LEBRD die 60° - und die 300° -KBRD. Damit werden die 0° -KBRD sowohl direkt über Linking-Verbindungen von den 0° -LEBRD als auch indirekt durch Inhibitor der 300° -KBRD unterstützt. Die Abbildung 3.94 B zeigt die starke Aktivität der 0° -KBRD als Ergebnis dieser Interaktion.

Aufgrund der isotropen Verschaltung der LEBRD zu den KBRD darf die Inhibition nicht auf die verbleibenden drei KBRD-Schichten ausgedehnt werden. Objekte, die sich außerhalb des gewöhnlichen „Barberpole Stimulus“ befinden und sich z.B. in die entgegengesetzte Richtung bewegen, könnten dann eventuell nicht mehr richtig detektiert werden. Die Detektion solcher Bewegungskontraste scheint aber gerade für die Objektseparation wichtig zu sein (vgl. Kapitel 2.1.3).

Abb. 3.93: Stimulus „Barberpole Illusion“. Ein Gitter mit 30° orientierten Balken wird hinter einer Verdeckung mit einer rechteckigen Apertur in Richtung 300° bewegt. Menschliche Beobachter nehmen hierbei eine Bewegung in Richtung 0° wahr. Diese Illusion könnte auf dem Einfluß der Terminatorbewegung auf die Bewegungsdetektion beruhen. Zur Veranschaulichung ist die Apertur gestrichelt dargestellt und die Terminatoren sind durch Kreise markiert.



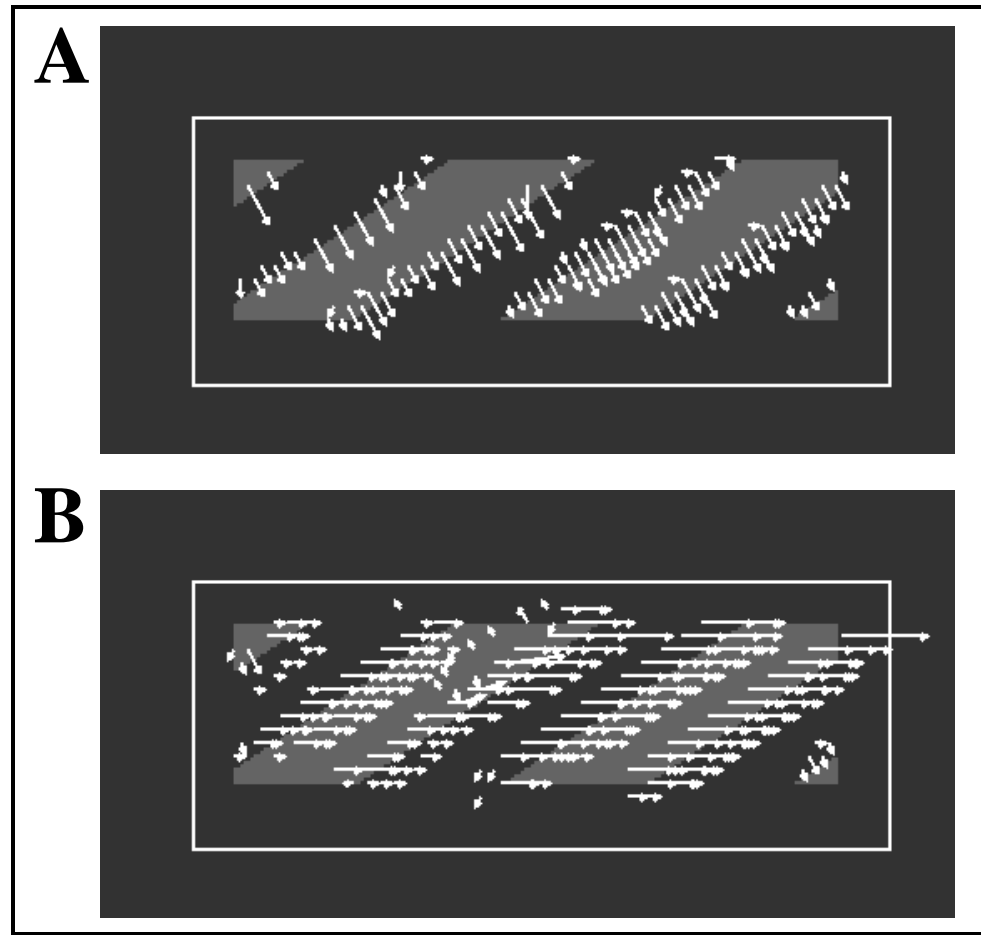


Abb. 3.94: Simulationsergebnis „Barberpole Illusion“. Dargestellt ist die Aktivität der ON- und OFF-Richtungsdetektoren für Konturen ohne (A) und mit (B) Interaktion zwischen Richtungsdetektoren für Linienenden und Richtungsdetektoren für Konturen. Das Simulationsergebnis in B spiegelt die menschliche Wahrnehmung bei der *barberpole illusion* wider. (Die Terminatorgeschwindigkeit in Richtung 0° betrug 8 Pixel/Bild mit $\Delta = 1$ Pixel.)

In der Abbildung 3.95 sind die über die gesamte Simulationsdauer ($t \geq 50$ Bin) kumulierten Aktivitäten der ON- und OFF-Richtungsdetektoren für vier verschiedene ($7 \text{ Pixel/Bild} \leq v_{0^\circ} \leq 10 \text{ Pixel/Bild}$) Terminatorgeschwindigkeiten dargestellt. Die aufgetragenen Werte sind über vier Positionsverschiebungen sowie über ON- und OFF-Pfad gemittelte Aktivitäten. Ohne Interaktion sind bei allen vier Geschwindigkeiten die 300° -Richtungsdetektoren, mit Interaktion hingegen die 0° -Richtungsdetektoren am aktivsten. Am deutlichsten zeigt sich die Wirkung der Interaktion bei einer Linienendengeschwindigkeit von $v_{0^\circ} = 8 \text{ Pixel/Bild}$, da bei dieser Geschwindigkeit die Richtungsdetektoren für Linienenden am besten stimuliert werden. Je größer die Abweichung der Terminatorgeschwindigkeit von der Geschwindigkeitspräferenz der LEBRD ist, desto schwächer sind natürlich auch die Interaktion und der Effekt (siehe Abbildung 3.95, $v_{0^\circ} = 10 \text{ Pixel/Bild}$). Dieses Problem kann aber einfach durch zusätzliche Richtungsdetektoren für Linienenden mit anderen Geschwindigkeitspräferenzen behoben werden. Allerdings müssen dafür viele Modellneuronen zur Verfügung gestellt werden, die in der verfügbaren Entwicklungsumgebung nicht vorhanden sind.

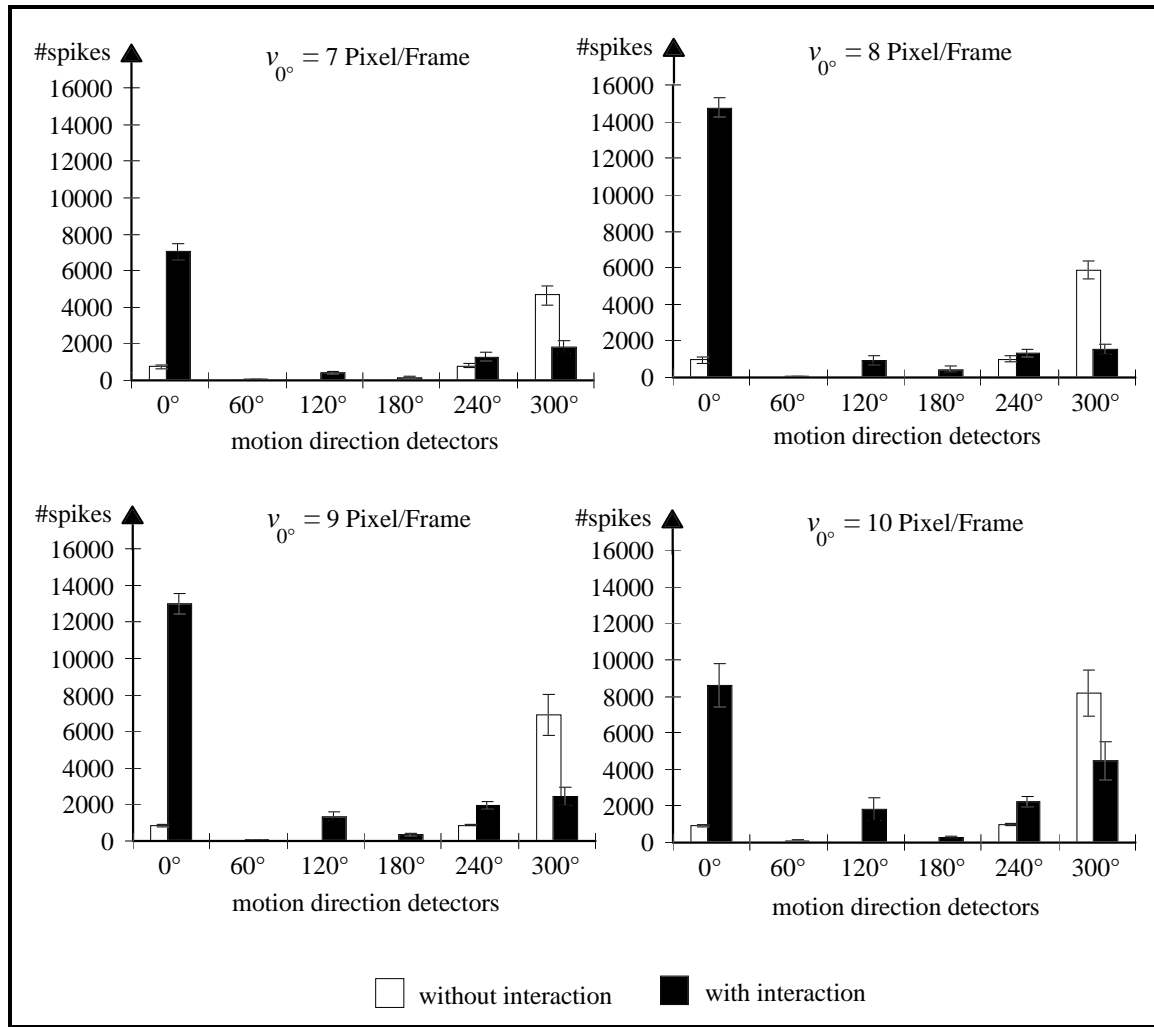


Abb. 3.95: Kumulierte Aktivitäten der Richtungsdetektoren für Konturen bei der „Barberpole Illusion“. Für vier verschiedene Geschwindigkeiten der Terminatoren in Richtung 0° sind die gemittelten Aktivitäten der ON- und OFF-Richtungsdetektoren aufgetragen. Ohne Interaktion der Richtungsdetektoren für Linienenden mit den Richtungsdetektoren für Konturen dominieren die 300° -Richtungsdetektoren, mit Interaktion hingegen die 0° -Richtungsdetektoren. (Stimulus wie in Abbildung 3.93.)

Bei dem „Barberpole-Effekt“ ist auch der in Abbildung 3.96 für eine Simulation dargestellte zeitliche Verlauf der Aktivitäten der Richtungsdetektoren für Konturen interessant. Hierfür wurden die von den 0° - und den 300° -Richtungsdetektoren pro Bild erzeugten Aktionspotentiale gemäß Gleichung 3.86 berechnet und gegen die Zeit aufgetragen. In den ersten 32 Bin der Simulation, dargestellt zum Zeitschritt $t = 0$ Bin, ist noch kein Richtungsdetektor aktiv. Danach kommt es zu einem deutlichen Anstieg in der Aktivität der 300° -Richtungsdetektoren, die bis zum Zeitschritt $t = 192$ Bin auf ihr minimales Niveau abfällt. Die 0° -Richtungsdetektoren sind in den ersten 100 Simulationsschritten kaum aktiv. Ihre Aktivität steigt danach aber stark an. Dieser verzögerte Anstieg liegt in der notwendigen Dauer, die Linienenden und deren Bewegungsrichtung in einer ausreichend starken Aktivität zu detektieren. Die Stärke des Anstiegs resultiert aus der zusätzlich wirkenden lateralen Kopplung der Richtungsdetektoren.

$$Y_{0^\circ, \text{ON}}(n \cdot T) = \sum_{t=n \cdot T}^{n \cdot T + 31 \text{ Bin}} \text{MUA}_{0^\circ, \text{ON}}(t) \quad \text{mit } 0 \leq n \leq 16, T = 32 \text{ Bin} \quad (\text{Gl. 3.86})$$

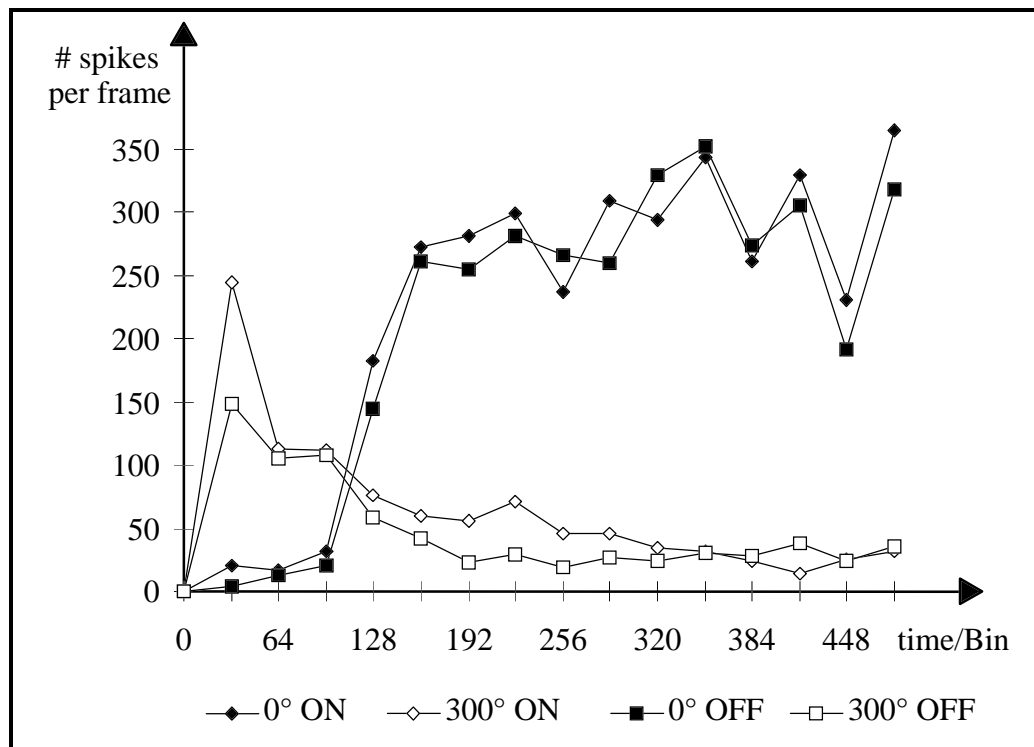


Abb. 3.96: Zeitlicher Verlauf der Aktivitäten der Richtungsdetektoren bei der Simulation der „Barberpole Illusion“. Dargestellt ist die jeweils über 32 Bin kumulierte Aktivität der 0°- und 300°-Richtungsdetektoren für Konturen gegen die Zeit. Die Aktivität der 0°-Richtungsdetektoren ist erst nach ungefähr 100 Simulationsschritten größer als die der 300°-Richtungsdetektoren. Dies ist auf die Dauer zurückzuführen, die das neuronale Netz braucht, um die Bewegungsrichtung der Linienenden zu bestimmen und zur Interaktion der Richtungsdetektoren für Linienenden mit den Richtungsdetektoren für Konturen.

Keine Barberpole Illusion mit eingeschnittener Apertur

Nach Kooi 1993 wird die *barberpole illusion* zerstört, wenn die Apertur eingeschnitten wird. Dies ist ein starker Hinweis darauf, daß die *barberpole illusion* von den lokalen Terminatorbewegungen und nicht nur von der Aperturausdehnung abhängt. In der nächsten Simulation wird deshalb die Apertur so eingeschnitten, daß sich die Terminatoren nicht mehr in Richtung 0°, sondern in Richtung 315° bewegen (Abbildung 3.97).

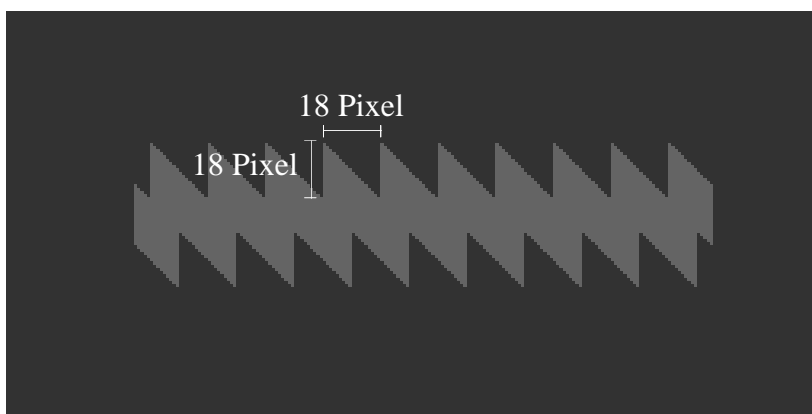


Abb. 3.97: Eingeschnittene Apertur. Die Abbildung zeigt die von Kooi 1993 beschriebene Apertur (hellgrau) zur Störung der *barberpole illusion*. Die Verdeckung wurde so eingeschnitten, daß die Terminatoren entlang der längeren Ausdehnung sich nicht mehr in Richtung 0°, sondern in Richtung 315° bewegen.

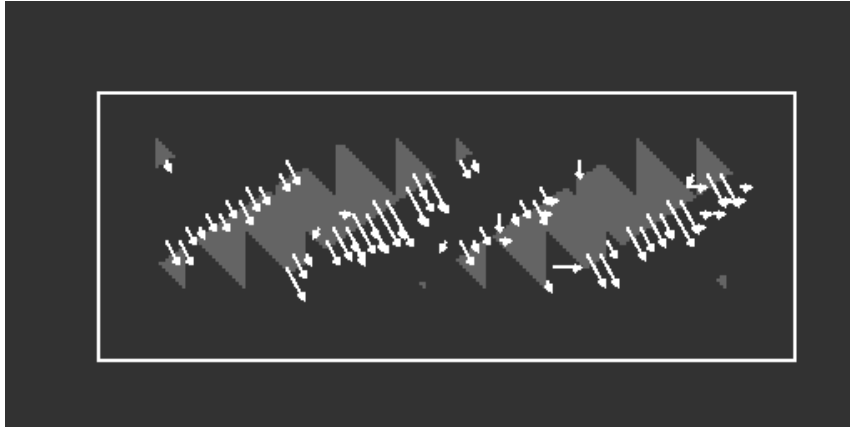


Abb. 3.98: Simulationsergebnis mit eingeschnittener Apertur. Dargestellt ist die Aktivität der ON- und OFF-Richtungsdetektoren für Konturen. Hauptsächlich sind die Modellneuronen mit einer Richtungspräferenz von 300° aktiv. (Die Geschwindigkeit der Terminatoren in Richtung 0° betrug 10 Pixel/Bild mit $\Delta = 1$ Pixel.)

Das Ergebnis dieser Simulation ist in den Abbildungen 3.98 und 3.99 aufgetragen. Erwartungsgemäß sind die 300° - und nicht die 0° -Richtungsdetektoren für Konturen am aktivsten. Da keine 0° -Bewegungsrichtung der Terminatoren mehr vorliegt, werden die 300° -Richtungsdetektoren nicht mehr von den 0° -LEBRD gehemmt und gewinnen den lokalen Wettstreit gegen die 240° - und 0° -Richtungsdetektoren. Die Abbildung 3.99 zeigt eine etwa halb so große Aktivität der Richtungsdetektoren für Konturen als bei der Simulation ohne Einschnitte und ohne Interaktion. Dies ist zum einen auf die Einschnitte zurückzuführen, die die Kanten der sichtbaren Balken im Eingabebild zerstückeln, was zur Folge hat, daß sich die Richtungsdetektoren, deren receptive Felder entlang dieser Kanten liegen, koppeln können. Zum anderen gibt es wegen der neu hinzugekommenen Verdeckung auch weniger 30° orientierte Kanten im Eingabebild.

Aus Abbildung 3.99 wird auch ersichtlich, daß mit zunehmender Terminatorgeschwindigkeit v_{0° die Aktivität der 300° -Richtungsdetektoren größer wird. Bei $v_{0^\circ} = 10$ Pixel/Bild haben die Terminatoren eine Geschwindigkeit in Richtung 45° von $v_{45^\circ} \cong 6$ Pixel/Bild und kommen somit in den bevorzugten Geschwindigkeitsbereich der Richtungsdetektoren für Linienenden. Diese können somit die Richtungsdetektoren für Konturen beeinflussen. Aufgrund der kurzen Laufzeit der Terminatoren auf den Einschnitten kann allerdings der Voraktivierungsmechanismus der Richtungsdetektoren für Linienenden nicht ausreichend wirksam werden. Dieser Nachteil kann durch eine größere Apertur behoben werden, wodurch aber der Verarbeitungsausschnitt vergrößert werden muß, was wiederum eine größere Anzahl an Modellneuronen zur Folge hätte.

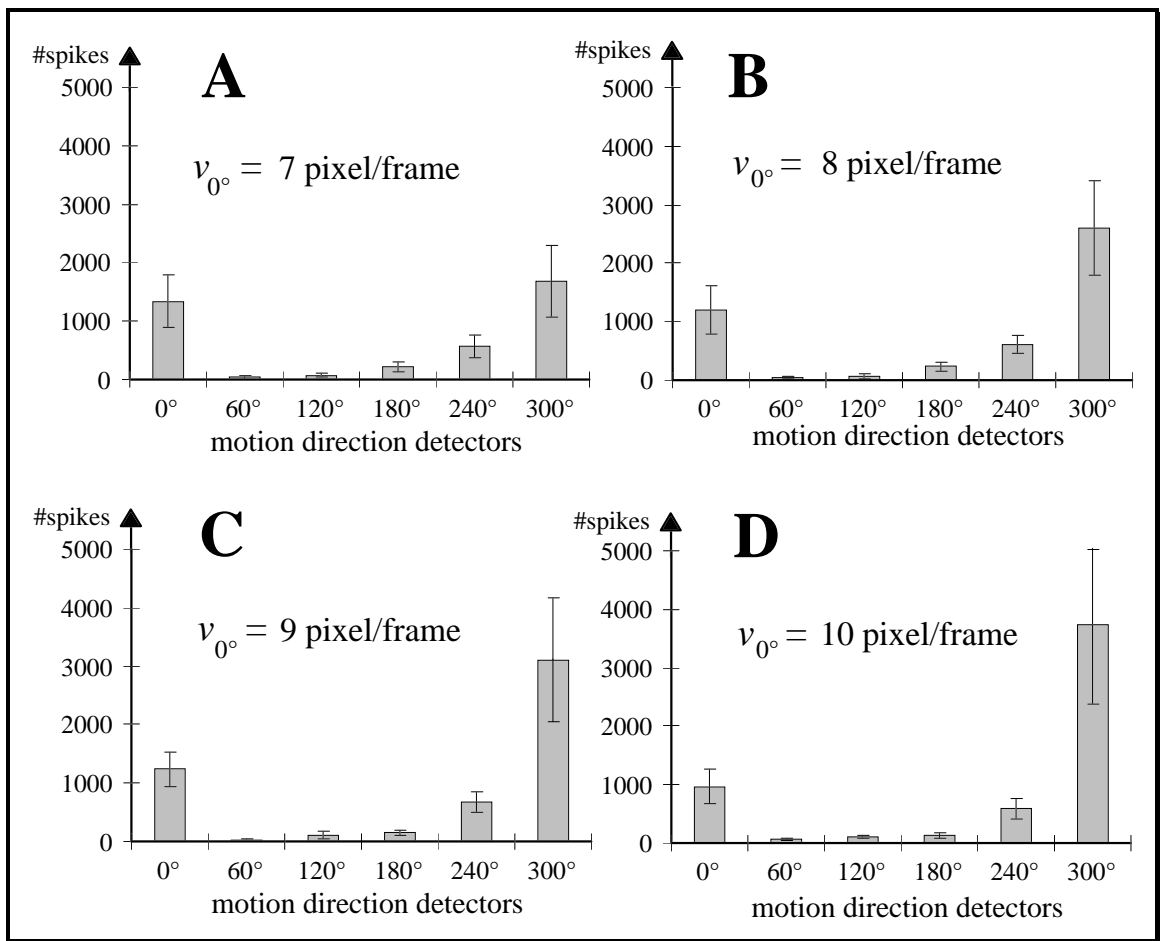


Abb. 3.99: Kumulierte Aktivitäten der Richtungsdetektoren für Konturen bei eingeschnittener Apertur. Für vier Terminatorgeschwindigkeiten in Richtung 0° sind die gemittelten Aktivitäten der ON- und OFF-Richtungsdetektoren für Konturen aufgetragen. Die 300° -Richtungsdetektoren sind am aktivsten. Das Simulationsergebnis ist konform mit den Untersuchungen Koois (1993), der zeigen konnte, daß keine *barberpole illusion* bei eingeschnittener Apertur auftritt.

Barberplaid Illusion

Mit dem beschriebenen Modell kann nicht nur die *barberpole illusion*, sondern auch die *barberplaid illusion* erklärt werden. Zur Simulation der *barberplaid illusion* wurden als Stimulus zwei Gitter hinter einer Verdeckung mit einer runden Apertur bewegt (Abbildung 3.100). Die Gitter bestanden aus Balken, deren Breite 36 Pixel war. Die Grauwerte der Balken betrugen $g_1 = 100$ und $g_2 = 0$. Die Balken mit Grauwert g_2 waren hierbei transparent, so daß durch die Überlagerung der beiden Gitter der typische „Barberplaid-Stimulus“ entstand. Die Gitterorientierungen waren 30° und 150° . Beide Gitter wurden mit einer Geschwindigkeitskomponente von $v = 8$ Pixel/Bild in Richtung 0° bewegt. Die Verdeckung hatte einen Grauwert von $g = 50$ und einen Aperturradius von 60 Pixel.

Werden diese Gitter einzeln durch die runde Apertur beobachtet, dann nehmen menschliche Beobachter eine Gitterbewegungsrichtung orthogonal zur Balkenorientierung - das wären in diesem Beispiel 300° für das 30° orientierte Gitter respektive 60° für das Gitter mit 150° orientierten Balken - wahr. Bei Bewegung beider Gitter entsteht hingegen ein Bewegungseindruck in Richtung 0° , wobei die Gitter nicht mehr getrennt, sondern als ein Objekt

wahrgenommen werden. Diese Wahrnehmung kann, wie bereits in Kapitel 2.2 geschildert wurde, ebenfalls auf die Bewegung der Terminatoren und ihren Einfluß auf die gesamte Bewegungsverarbeitung zurückgeführt werden.

Zur Simulation dieses psychophysischen Effektes wurde das gleiche neuronale Netz wie zuvor benutzt. Lediglich die Anzahl der Abtastorte wurde der Stimulusgröße angepaßt, so daß jede neuronale Schicht aus 32×32 Modellneuronen bestand. Für die Extraktion der Kantenorientierungen wurden nur die ON-Orientierungsdetektoren verwendet. Die Bewegungsrichtungen der Linienenden wurde mit $d = 8$ Pixel bestimmt. Die Simulationsdauer betrug $t = 1568$ Bin.

In der Abbildung 3.101 sind für ein Bild der Eingabesequenz die Aktivitäten der ON- und OFF-Richtungsdetektoren für Konturen mit (Abb. 3.101 A) und ohne (Abb. 3.101 B) Interaktion mit den LEBRD dargestellt. Die Abbildung 3.102 zeigt die über die gesamte Simulationsdauer kumulierten Aktivitäten der KBRD ohne den Einschalteffekt ($t < 50$ Bin). Die dargestellten Werte sind gemittelte Aktivitäten der ON- und OFF-Richtungsdetektoren für vier verschiedene Gitterpositionsverschiebungen.

Abb. 3.100: Stimulus „Barberplaid Illusion“. Zwei Gitter mit unterschiedlichen Orientierungen werden überlagert und hinter einer Verdeckung mit einer runden Apertur bewegt (z.B. Gitter I in Richtung 300° und Gitter II in Richtung 60°). Menschliche Beobachter nehmen bei ähnlicher Gitterfrequenz und -geschwindigkeit nicht mehr zwei Gitter, die sich orthogonal zu ihren Balkenorientierungen bewegen, wahr. Es entsteht vielmehr der Eindruck eines einzigen sich bewegenden Objektes, dessen Bewegungsrichtung der durch die Überlagerung der beiden Gitter entstehenden Terminatoren entspricht.

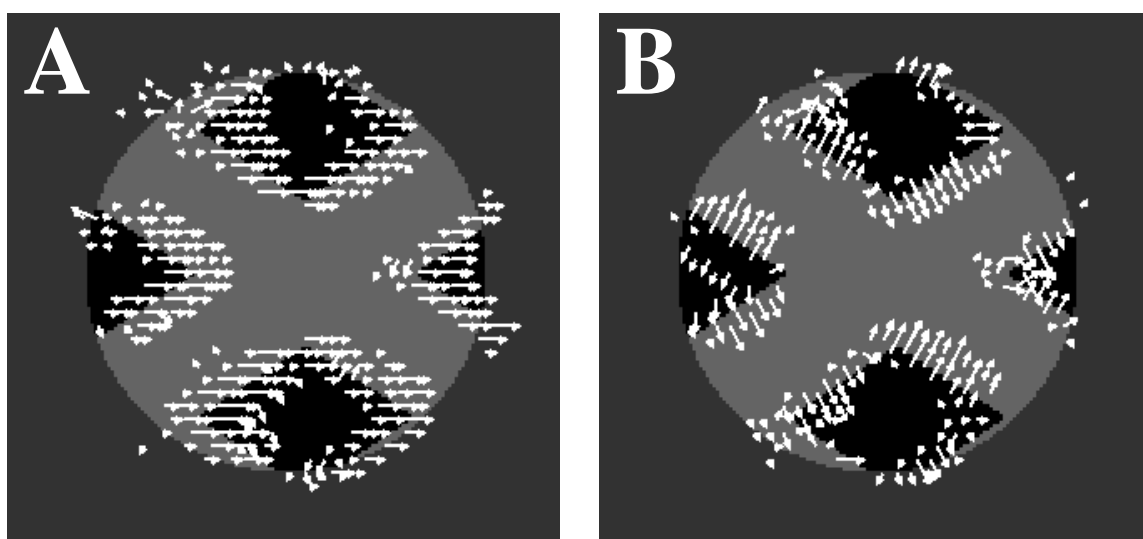
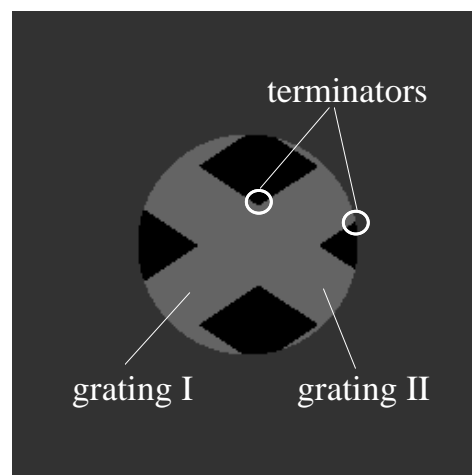


Abb. 3.101: Simulationsergebnis „Barberplaid Illusion“ I. Dargestellt sind die Aktivitäten der ON- und OFF-Richtungsdetektoren für Konturen als Pfeillängen. Die Simulation A wurde mit Interaktion zwischen den LEBRD und den KBRD durchgeführt und die Simulation B ohne Interaktion. In A dominieren KBRD mit einer Richtungspräferenz von 0° , in B die 60° und 300° Detektoren.

Ohne explizite Miteinbeziehung der Terminatorbewegungsrichtung sind die Detektoren mit Richtungspräferenzen von 60° und 300° am aktivsten. Die beiden Gitter werden in getrennten Richtungsdetektorschichten repräsentiert und können somit als separate Objekte angesehen werden. Die geringen Aktivitäten der 0° -, 120° - und 240° -Richtungsdetektoren sind hauptsächlich in Randeffekten begründet.

Wird hingegen die Bewegungsrichtungsinformation der Terminatoren extrahiert und über die oben beschriebene neuronale Verschaltung an die Richtungsdetektoren für Konturen weitergegeben, dann werden beide Gitter zusammen in der 0° -KBRD-Schicht repräsentiert. Die beiden Gitter sind durch die Richtungsdetektoren nicht mehr separierbar, sondern werden zu einem Objekt zusammengebunden.

In Abbildung 3.102 ist zur Kontrolle noch die Aktivität der ON- und OFF-Richtungsdetektoren bei Stimulation mit dem 150° orientierten Gitter aufgetragen. Auch mit Interaktion zwischen den Richtungsdetektoren für Konturen und den Richtungsdetektoren für Linienenden sind hauptsächlich die 60° -KBRD aktiv. Die nicht zu vernachlässigende Aktivierung der 0° - und 120° -KBRD resultiert aus der Einwirkung der Terminatorbewegungen am Rand der Apertur. Da die Anzahl der Terminatoren allerdings gering ist, ist auch deren Einfluß auf die Bewegungsverarbeitung begrenzt.

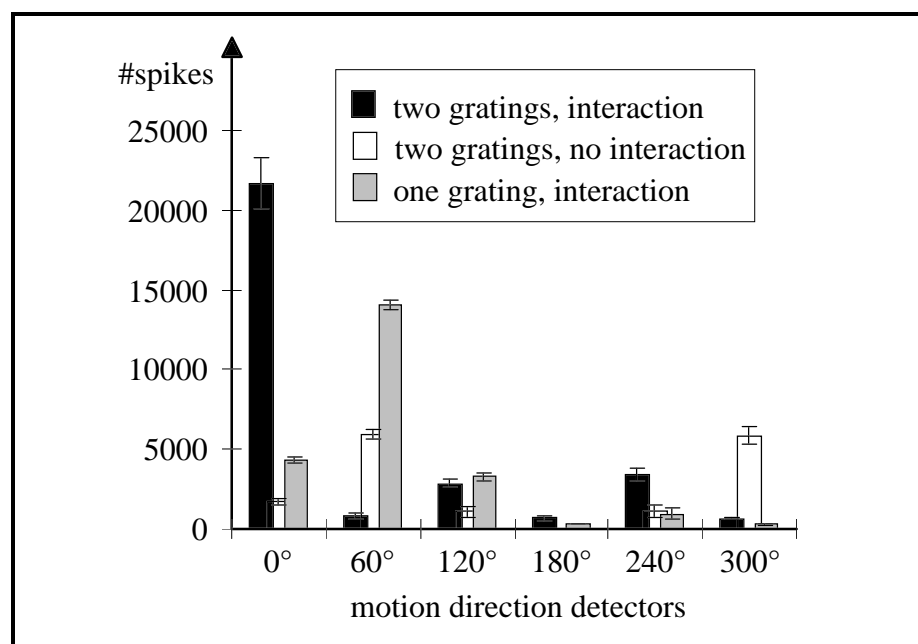


Abb. 3.102: Simulationsergebnis „Barberplaid Illusion“ II. Dargestellt sind die kumulierten Aktivitäten der ON- und OFF-Richtungsdetektoren für drei unterschiedliche Simulationsbedingungen. Die schwarzen und weißen Balken zeigen die Aktivitäten bei Stimulation mit zwei Gittern (30° und 150° orientiert und in Richtung 0° mit $v = 8$ Pixel/Bild bewegt) mit bzw. ohne Einwirkung der LEBRD auf die KBRD. Mit Interaktion sind hauptsächlich die 0° -Richtungsdetektoren für Konturen aktiv, da sich die durch die Überschneidungen der Gitter ergebenden Terminatoren in Richtung 0° bewegen. Ohne Interaktion dominieren die Bewegungsrichtungen orthogonal zu den Gitterorientierungen. Die grauen Balken zeigen zur Kontrolle die Aktivität bei Stimulation mit dem 150° orientierten Gitter alleine.

3.4.2.2 Linienendendetektion in realen Szenen

Im letzten Kapitel wurde der Einfluß der Linienenden auf die Bewegungsdetektion anhand einer Simulation mit künstlicher Eingabe gezeigt. Hierbei konnte eine mögliche Erklärung für die *barberpole illusion* gegeben werden. In diesem Kapitel soll hingegen die Verwertbarkeit der Linienenden in der Bewegungsdetektion bei realen Szenen untersucht werden. Es wird sich zeigen, daß Linienenden zwar aus den Eingabebildern extrahiert werden können, doch daß dessen Bewegungsrichtungsbestimmung schwierig ist.

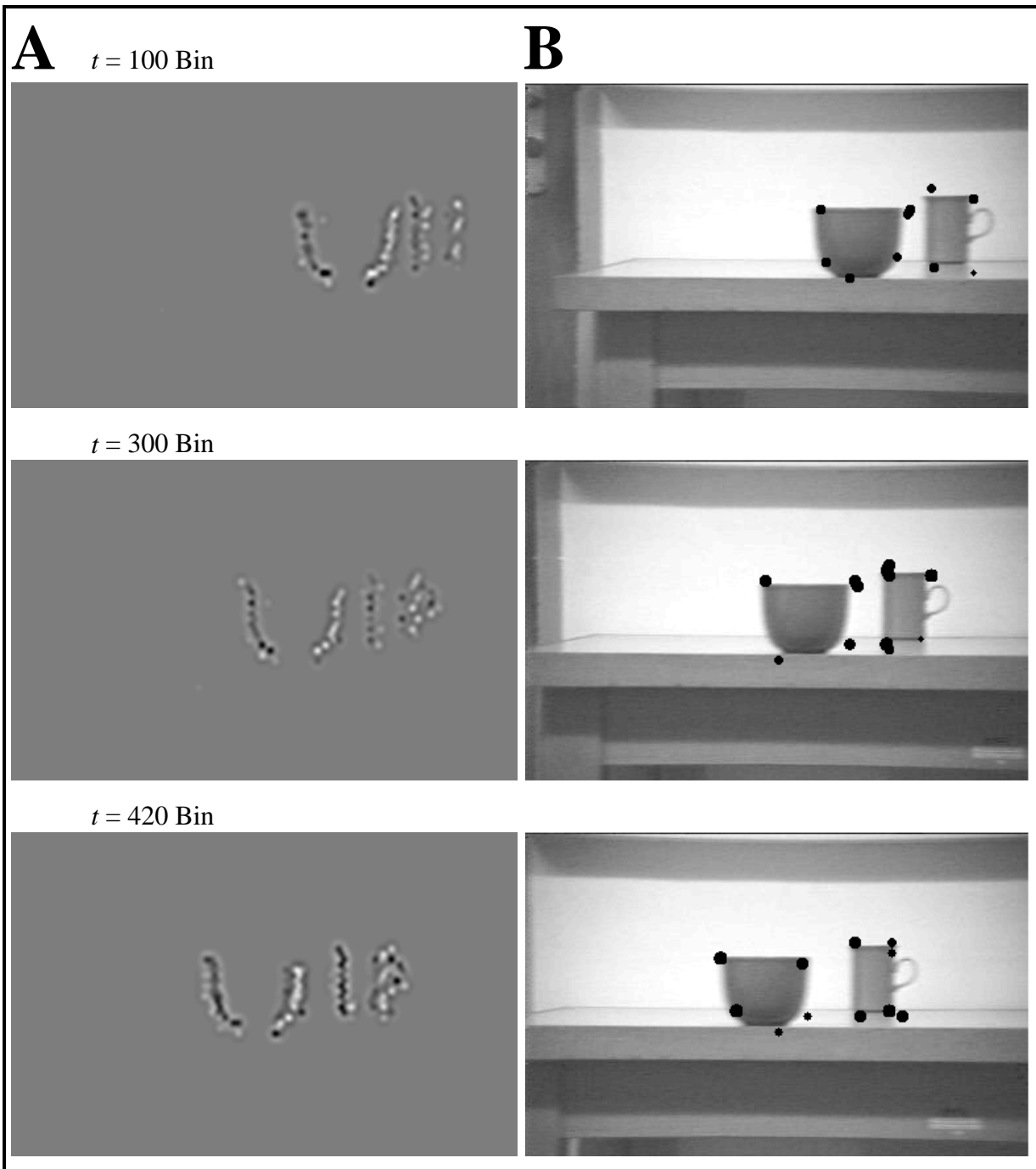


Abb. 3.103: Linienendendetektion „Tasse“. **A:** Rückprojektion der rezeptiven Felder aktiver retinaler Ganglienzellen in den Bildbereich. **B:** Markierung der vom System extrahierten Linienenden. Erläuterungen im Text.

Bei der Aufnahme der ersten zu verarbeitenden Sequenz wurde eine Kamera an zwei sich in unterschiedlicher Entfernung zur Kamera befindenden Tassen vorbeigeführt. Dieses Beispiel soll noch einmal deutlich machen, daß im Transientensystem ausschließlich sich lokal verändernde Intensitätskontraste verarbeitet werden. Die Abbildung 3.103 A zeigt zu diesem Zweck die Rückprojektion aktiver retinaler Y-Ganglienzellen in den Bildbereich (Kopecz et al. 1996) zu drei verschiedenen Simulationszeiten. Es wurden Verarbeitungseinheiten der mittleren Auflösung ($d = 4$ Pixel) benutzt. Die parallel zur Bewegungsrichtung orientierten Intensitätskontrastkanten, im Falle der Tassen sind dies die obere und untere Kante, führen zu keiner Aktivierung retinaler Ganglienzellen. Der Henkel der rechten Tasse wird aufgrund seiner geringen Breite in den Eingabebildern von den retinalen Ganglienzellen der mittleren Auflösung nur unzureichend detektiert. Die Rückprojektion zeigt allerdings, daß die Linienenden der langen, transienten Intensitätskontrastkanten extrahierbar sind. Das Ergebnis dieser Linienendenextraktion wird in der Abbildung 3.103 B gezeigt. Aktive Linienendendetektoren sind entsprechend ihrer Lage im Abtastgitter durch schwarze Flecken im Bild dargestellt. Alle Ecken der Tassen im Bild werden, wenn auch nicht ständig, detektiert. Bei der linken Tasse sind zusätzlich Linienendendetektoren aktiv, die die starke Krümmung der Tasse im unteren Bereich detektieren.

In dieser Simulation zeigt sich aber auch schon ein Problem. Die Linienenden werden aufgrund der Bewegung und der Unterabtastung nicht immer an derselben Stelle des Objektes detektiert. Weiterhin kann ein Linienende zu einer Aktivierung von zwei Linienendendetektoren führen. Werden beide Linienendendetektoren gleich stark aktiviert, dann kann auch der lokale Wettstreit über das inhibitorisch wirkende Interneuron zu keiner eindeutigen Detektorantwort führen. Beide Probleme sind nachteilig für eine Bewegungsrichtungsdetektion der Linienenden.

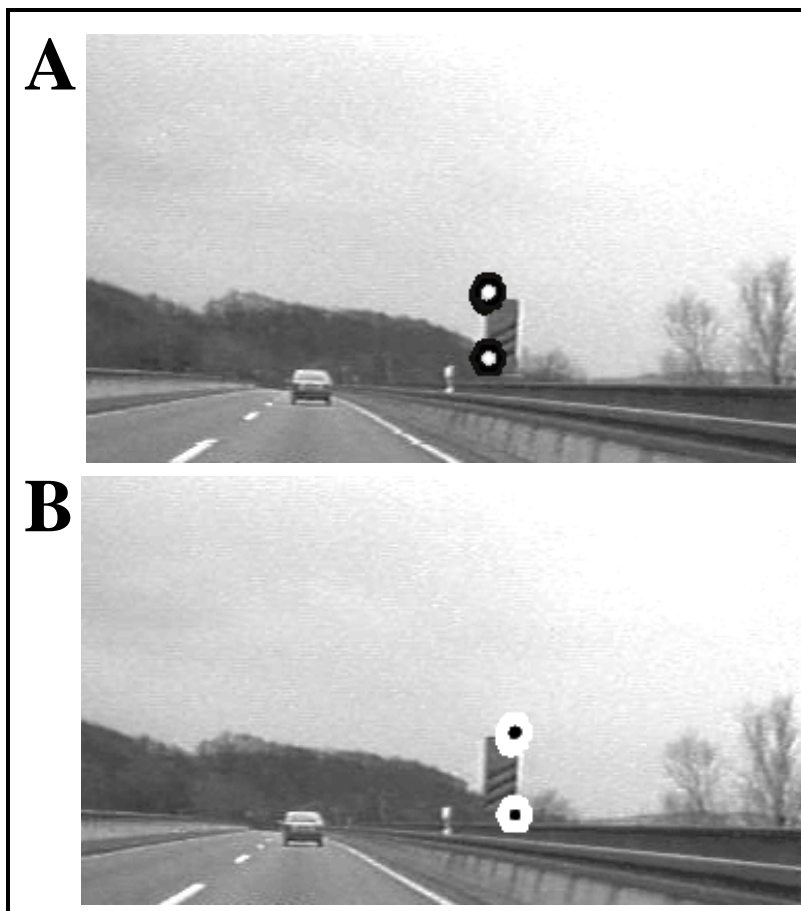


Abb. 3.104: Extraktion der Linienenden eines Verkehrsschildes. Die Abbildung zeigt in **A** die Extraktion der Linienenden im ON-Pfad und in **B** die der Linienenden im OFF-Pfad. Die aktiven Linienendendetektoren sind im Bild durch die Rückprojektionen der rezeptiven Felder der retinalen Ganglienzellen mit gleichem RF-Mittelpunkt dargestellt. Die Linienenden der linken Seite des Verkehrsschildes werden im ON-Pfad und die Linienenden der rechten Seite im OFF-Pfad extrahiert.

Eine weitere Simulation, deren Ergebnis in Abbildung 3.104 dargestellt ist, zeigt die Extraktion der Linienenden in einer Sequenz mit einem Verkehrsschild. Es wurden wiederum die Verarbeitungseinheiten der mittleren Auflösung verwendet. Das Beispiel soll verdeutlichen, daß die Linienenden in zwei unterschiedlichen Pfaden extrahiert werden. Dazu wurde in dieser Simulation auf die Zusammenlegung des ON- und OFF-Pfades verzichtet und beide Pfade separat simuliert. Auch in diesem Beispiel zeigt sich die Fähigkeit des Systems, Linienenden zu extrahieren. Das im Vergleich zum Hintergrund dunkle Schild hat im Eingabebild zwei sich orthogonal zur Bewegungsrichtung bewegende Intensitätskontrastkanten: auf der linken Seite des Schildes eine hell/dunkel-Kante und auf der rechten Seite eine dunkel/hell-Kante. Dadurch wird die linke Kante hauptsächlich im ON-Pfad verarbeitet, währenddessen die rechte Kante im OFF-Pfad verarbeitet wird. Demzufolge werden auch die Linienenden der linken Seite von Linienendendetektoren des ON-Pfades extrahiert und die der rechten Seite von Linienendendetektoren des OFF-Pfades.

In einem weiteren Beispiel wird die Sequenz „Gerade“ als Eingabe benutzt. Die Simulation wurde mit den Detektoren der feinen Auflösung durchgeführt. Die Abbildung 3.105 zeigt vier Bilder der Sequenz mit den von den Linienendendetektoren extrahierten Linienenden. Die Simulation wurde mit der Zusammenlegung von ON- und OFF-Pfad durchgeführt. Die Markierungen wurden aufgrund der besseren Sichtbarkeit in den Bildern innen schwarz und außen weiß dargestellt. Selbst in der feinen Auflösung wird das Kraftfahrzeug in dieser Entfernung nur noch durch fünf Linienenden beschrieben. Eine Klassifikation anhand der Linienenden ist damit eher als unwahrscheinlich anzusehen.



Abb. 3.105: Linienendendetektion in der Sequenz „Gerade“: Die Abbildung zeigt mittels der Markierungen (außen weiß, innen schwarz) in vier Bildern der Sequenz die Linienendendetektion (feine Auflösung). Die Simulation wurde mit der Zusammenlegung von ON- und OFF-Pfad durchgeführt. Somit stammen die detektierten Linienenden aus beiden Pfaden.

Die Bewegungsrichtungsbestimmung der Linienenden in der Sequenz „Gerade“ führt zu den in Abbildung 3.106 dargestellten Ergebnissen. Hierbei wurden ON- und OFF-Pfad getrennt verarbeitet. Die Antwort der Richtungsdetektoren für Linienenden ist eindeutig. Es wird eine Bewegung des Objektes im Bild nach rechts angezeigt. Allerdings wird die Bestimmung der Bewegungsrichtung nicht immer an allen Linienenden erfolgreich vorgenommen. Störungen in Form von Wackeln der Objektkanten, was zum Beispiel bei einer sich mitbewegenden Kamera im Straßenverkehr leicht vorkommen kann, behindern die Bewegungsrichtungsextraktion an der Kontur nahezu überhaupt nicht, wohingegen die Bewegungsrichtungsbestimmung der Linienenden aufgrund ihrer ungleichmäßigen Bewegung störanfällig ist. In einer technischen Anwendung sollte deshalb der Bewegungsrichtungsbestimmung an der Kontur der der Linienenden den Vorzug gegeben werden.



Abb. 3.106: Bestimmung der Linienendenbewegungsrichtungen. Durch die eingezeichneten Pfeile werden die Bewegungsrichtungen der Linienenden dargestellt (Kopecz et al. 1996). Die Länge der Pfeile gibt hierbei die Stärke der Aktivierung an. Die Verarbeitung erfolgte mit der feinen Auflösung in zwei getrennten Pfaden (A: ON und B: OFF).

3.5 Bewegungskontrast

Eine Objektdefinition aufgrund der absoluten Bewegungsrichtung ist nicht immer möglich und sinnvoll. Dies zeigte bereits das in Kapitel 2.1.3 beschriebene Beispiel der gegenläufig rotierenden Kreisscheiben. Die beiden Scheiben können nicht voneinander separiert werden, wenn ausschließlich die Bewegungsrichtung zur Gruppierung herangezogen wird, da bei beiden Scheiben alle Bewegungsrichtungen vorkommen. Menschliche Beobachter sind aber trotzdem in der Lage, die beiden Scheiben voneinander zu separieren. Eine genauere Betrachtung zeigt, daß auch beim Merkmal „Bewegung“ lokale Änderungen wichtig sind und zur Objektseparation herangezogen werden können (Nothdurft 1993, Nothdurft und Parlitz 1993, Lüschoff und Nothdurft 1993). Dies zeigt auch das Beispiel der Kreisscheiben. Innerhalb einer rotierenden Kreisscheibe ändert sich die Bewegungsrichtung benachbarter Punkte nur gering. An der Grenze zwischen den Kreisscheiben liegt hingegen eine große lokale Bewegungsrichtungsänderung vor (*Bewegungskontrast*). Somit liegt es nahe, eine Objektseparation durch eine lokale Bewegungskontrastdetektion durchzuführen. In Kapitel 3.5.1 wird dazu ein neuronales Netz vorgestellt, das lokalen Bewegungskontrast extrahieren kann (Fenske et al. 1995, Schott et al. 1996). Die Anwendbarkeit dieses neuronalen Netzes wird danach in Kapitel 3.5.2 anhand einiger Beispiele demonstriert. Das Prinzip der lokalen Kontrastdetektion ist auf das Merkmal „Orientierung“ übertragbar und kann zur Texturanalyse eingesetzt werden (Weitzel et al. 1995, Weitzel 1998).

3.5.1 Modell zur Extraktion des lokalen Bewegungskontrastes

Das neuronale Netzwerk zur Extraktion des lokalen Bewegungskontrastes schließt sich dem Modell zur Bestimmung der Bewegungsrichtung, das in Kapitel 3.2 vorgestellt wurde, an. In der nachfolgenden Abbildung 3.107 ist das Modell in einer Übersicht dargestellt.

Nach der Bestimmung der lokalen Bewegungsrichtung werden zunächst die Antworten der Richtungsdetektoren mit räumlich benachbarten rezeptiven Feldern von Interneuronen zusammengefaßt (Abbildung 3.108). Die Interneuronen sind hierbei auf einem Abtastgitter mit der doppelten Abtastweite der Richtungsdetektorschichten angeordnet, wodurch sich sowohl die Modellneuronenanzahl als auch die Verbindungsanzahl reduziert. Hierbei bietet sich zusätzlich noch eine Zusammenführung von ON- und OFF-Pfad an. Die Verschaltung der Richtungsdetektoren zu den Interneuronen wird durch Gleichung 3.87 beschrieben. Durch die konvergente Verschaltung werden die Richtungsdetektorantworten räumlich geglättet, wodurch Bewegungsfelder entstehen, die über die nachfolgende Bewegungskontrastdetektion voneinander abgegrenzt werden.

$$\text{ON/OFF-M,}\varphi_{\text{feed1}} w_{ijkl}^{\text{M-Smooth,}\varphi} = \begin{cases} \text{M-Smooth } V_n^{\text{FI}}, & \text{für hexagon}_{2d}n \\ 0, & \text{sonst} \end{cases} \quad 0 \leq n \leq 2 \quad (\text{Gl. 3.87})$$

mit $\varphi \in \{0^\circ, 60^\circ, 120^\circ, 180^\circ, 240^\circ, 300^\circ\}$

Die Bewegungskontrastdetektion kann auf zwei Arten erfolgen. In Anlehnung an die Bestimmung des Intensitätskontrastes kann der Bewegungskontrast über eine räumliche Bandpaßfilterung der geglätteten Richtungsdetektoraktivität ermittelt werden. Die räumliche Bandpaßfilterung kann durch eine Faltung der geglätteten Richtungsdetektoraktivität mit einer Mexikanerhut-Funktion erfolgen. In der neuronalen Verschaltung wird dies durch eine Zentrum/Umfeld-Struktur bewerkstelligt. Das Zentrum wird hierbei positiv und das Umfeld negativ gewichtet. Ein homogenes Bewegungsfeld führt somit zu einem unterschwelligen Membranpotential. An den Rändern dieses Bewegungsfeldes entsteht hingegen ein überschwelliges Membranpotential, wodurch die Grenzen des Bewegungsfeldes durch aktive

Modellneuronen gekennzeichnet werden. Da kleine Änderungen in der Bewegungsrichtung nicht detektiert werden sollen, wird das Umfeld der neuronalen Verschaltung auf benachbarte Richtungspräferenzen ausgedehnt. Die mathematische Beschreibung der Verschaltung wird durch die Gleichungen 3.88 und 3.89 wiedergegeben. In Abbildung 3.109 ist sie graphisch dargestellt.

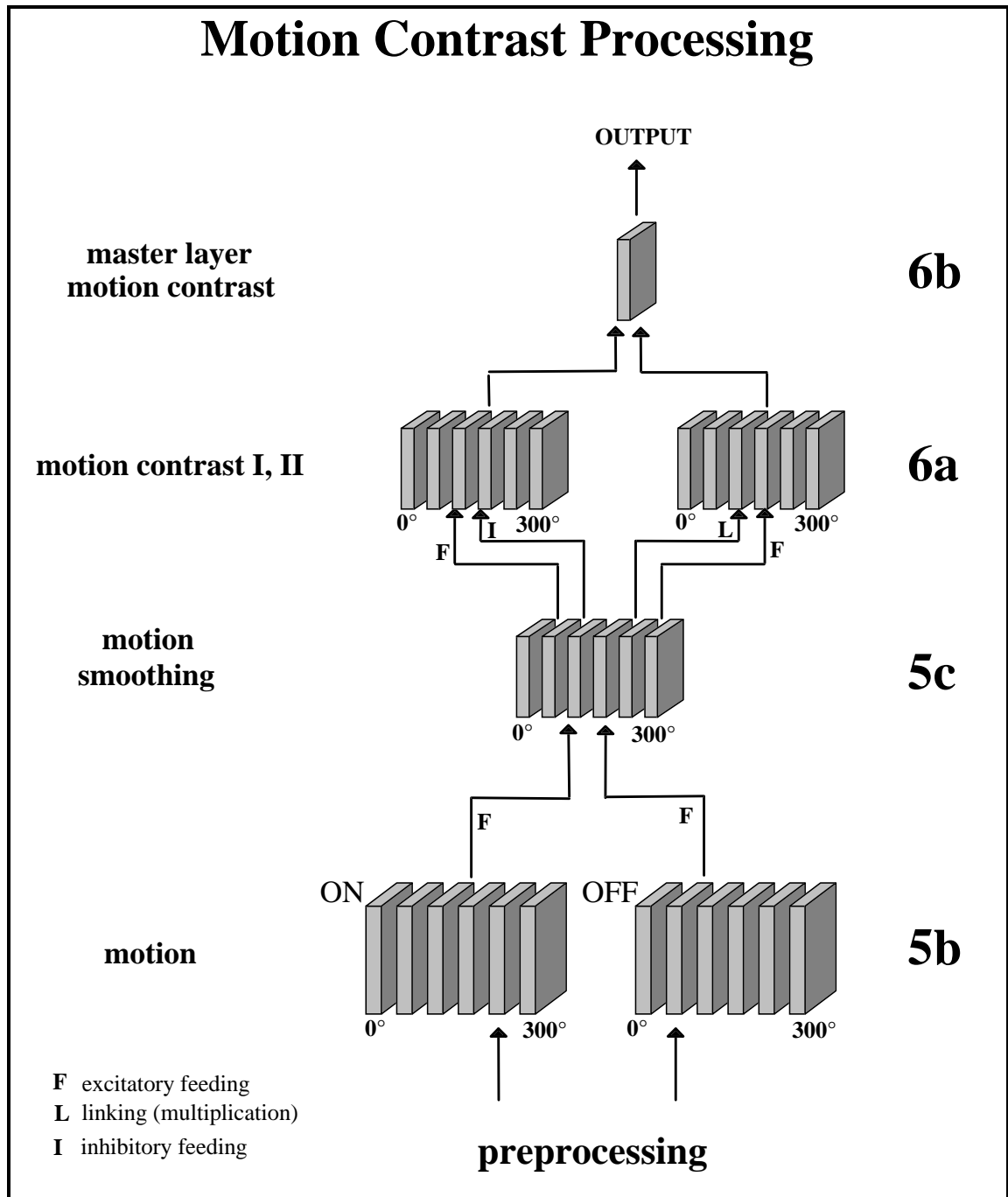


Abb. 3.107: Gesamtmodell zur Extraktion des Bewegungskontrastes. Das Modell zur Bestimmung des lokalen Bewegungskontrastes baut auf die Richtungsdetektion (5b) auf. Zur Reduzierung der Verbindungs- und Modellneuronenanzahl sowie zur räumlichen Glättung der Richtungsdetektoraktivität werden Interneuronen (5c) eingeführt. Anschließend kann auf zwei unterschiedliche Arten der lokale Bewegungskontrast detektiert werden (6a). Das Ergebnis dieser Detektion wird in einer Master-Bewegungskontrastschicht zusammengefaßt (6b).

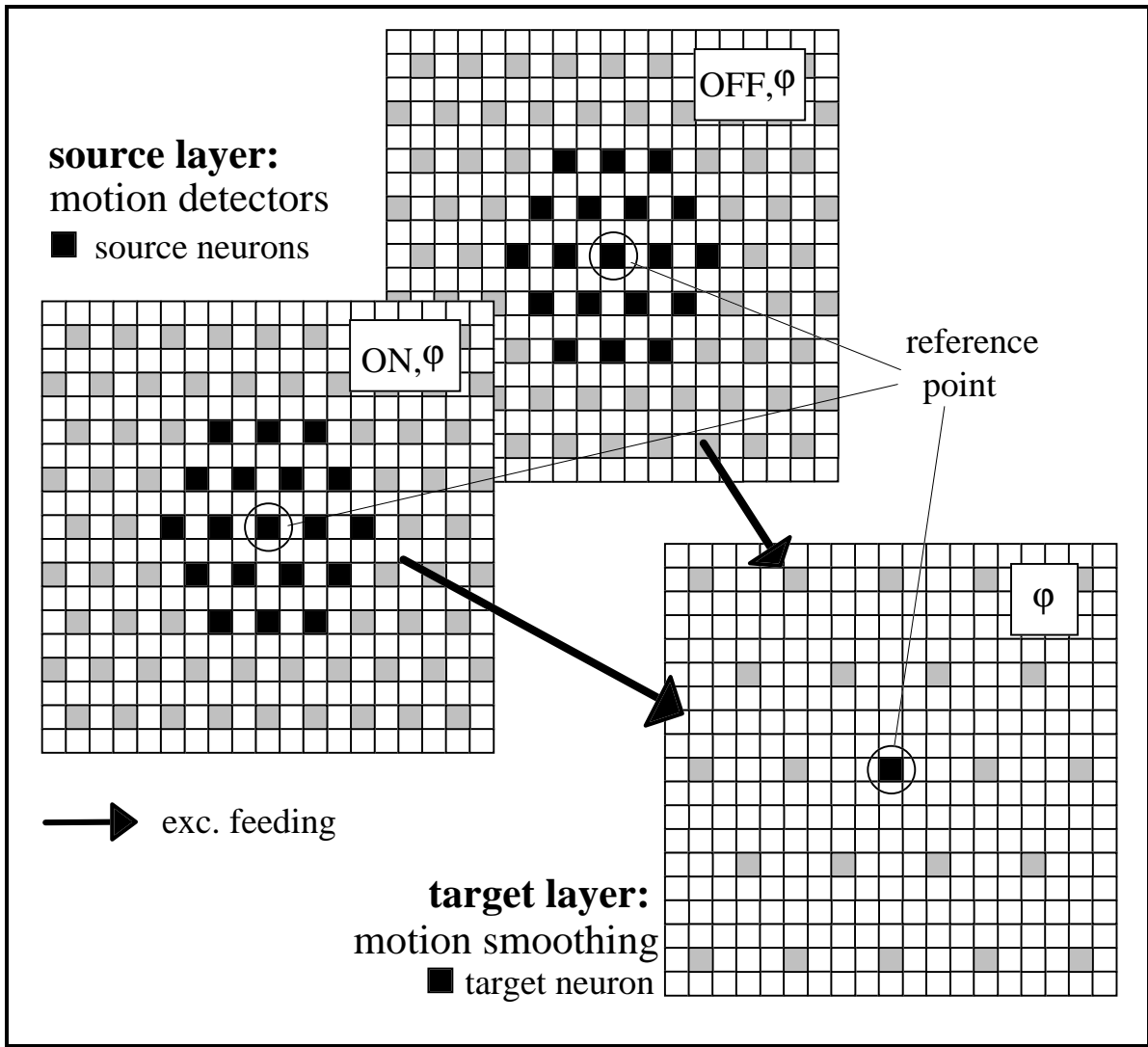


Abb. 3.108: Räumliche Glättung der Richtungsdetektoraktivität. Die Ausgänge der ON- und OFF-Richtungsdetektoren werden konvergent auf Interneuronen geschaltet, wodurch die Richtungsdetektoraktivität räumlich geglättet wird. Für jede Bewegungsrichtungspräferenz steht eine Schicht Interneuronen zur Verfügung. Die Abtastweite des Interneuronenabtastgitters ist doppelt so groß wie die der Richtungsdetektoren.

$$M\text{-Smooth}_{\text{feed1}}^{\varphi} w_{ijkl}^{\text{MC-I},\varphi} = \begin{cases} \text{MC-I } V^{\text{F1}}, & \text{falls } ((i = k) \wedge (j = l)) \\ 0, & \text{sonst} \end{cases} \quad (\text{Gl. 3.88})$$

$$M\text{-Smooth}_{\text{inh}}^{\varphi-60^\circ \leq \varphi' \leq \varphi+60^\circ} w_{ijkl}^{\text{MC-I},\varphi} = \begin{cases} \text{MC-I } V^{\text{I}}, & \text{für hexagon}_{2d} 1 \\ 0, & \text{sonst} \end{cases} \quad (\text{Gl. 3.89})$$

mit $\varphi \in \{0^\circ, 60^\circ, 120^\circ, 180^\circ, 240^\circ, 300^\circ\}$

Eine zweite Möglichkeit, den lokalen Bewegungskontrast zu bestimmen, ist die lokale Verknüpfung stark unterschiedlicher Bewegungsrichtungen. Tritt in einem bestimmten Umfeld einer von einem Objekt verursachten Bewegungsrichtung eine weitere, unähnliche Bewegungsrichtung auf, dann gehört die zweite Bewegungsrichtung wahrscheinlich zu einem anderen Objekt, was durch aktive Bewegungskontrastdetektoren angezeigt werden soll. Hierbei zeichnet sich die Verschaltung ebenfalls durch eine Zentrum/Umfeld-Struktur aus. Im Unterschied zu der in den Gleichungen 3.88 und 3.89 beschriebenen Implementierung wird das

Umfeld über multiplikativ wirkende Linking-Verbindungen auf den Bewegungskontrastdetektor geschaltet. Außerdem entstammt das Umfeld nicht den Richtungsdetektorschichten ähnlicher Richtungspräferenzen, sondern aus denen mit mindestens 120° Unterschied in der Richtungspräferenz. Die Gleichungen 3.90 und 3.91 beschreiben diese Verschaltungsstruktur.

$$M\text{-Smooth}_{\text{feed1}}^{\varphi} w_{ijkl}^{MC-II,\varphi} = \begin{cases} MC-II V^{F1}, & \text{falls } ((i = k) \wedge (j = l)) \\ 0, & \text{sonst} \end{cases} \quad (\text{Gl. 3.90})$$

$$M\text{-Smooth}_{\text{link,mult}}^{\varphi'} w_{ijkl}^{MC-II,\varphi} = \begin{cases} MC-II V^L, & \text{für hexagon}_{2d} 1 \\ 0, & \text{sonst} \end{cases} \quad (\text{Gl. 3.91})$$

mit $\varphi \in \{0^\circ, 60^\circ, 120^\circ, 180^\circ, 240^\circ, 300^\circ\}$

und $\varphi' = \varphi + 180^\circ$ oder $\varphi' = \varphi + 180^\circ \pm 60^\circ$

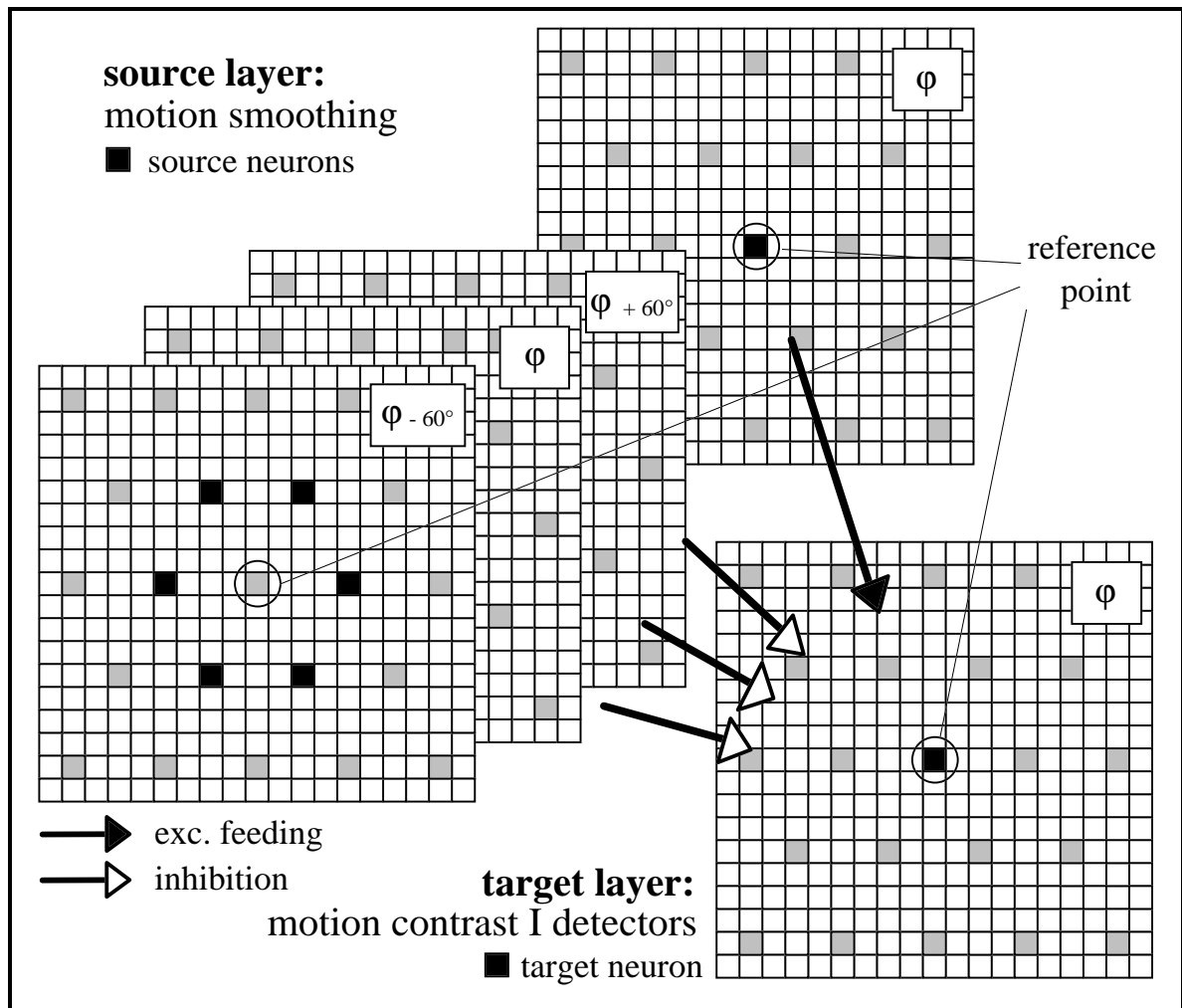


Abb. 3.109: Modell „Bewegungskontrast I“. Die Verschaltung zur Bestimmung des Bewegungskontrastes I hat eine Zentrum/Umfeld-Struktur. Das Zentrum wirkt hierbei exzitatorisch und das Umfeld inhibitorisch. In homogenen Bewegungsfeldern wird somit ein unterschwelliges Membranpotential erzeugt. Ränder des Bewegungsfeldes führen zu einem überschwelligen Membranpotential. Damit geringe Richtungsänderungen nicht detektiert werden, wird das Umfeld auf benachbarte Richtungspräferenzen ausgedehnt.

Beide Mechanismen zur Bewegungskontrastdetektion können entweder zusammen oder einzeln implementiert werden. Dies hängt wiederum von der Anzahl der im technischen System zur Verfügung stehenden Modellneuronen ab.

Die Aktivität der einzelnen Bewegungskontrastdetektorschichten wird anschließend in einer Master-Schicht nach Gleichung 3.92 zusammengefaßt. Der in ihr repräsentierte Bewegungskontrast kann dann zur Objektseparation (z.B. Pauly et al. 1996) oder zur Aufmerksamkeitssteuerung (z.B. Pauly et al. 1997) herangezogen werden. Denkbar wäre auch eine Zuführung der Bewegungskontrastdetektoraktivität zu Orientierungs- und Linienendendetektoren, um die Form des Objektes, die durch den Bewegungskontrast beschrieben wird, für eine Objekterkennung näher analysieren zu können. Die Parameter des Modelles zur Extraktion des Bewegungskontrastes sind in Tabelle 3.9 aufgeführt.

$${}_{\text{feed1}}^{\text{MC}, \varphi'} w_{ijkl}^{\text{Master-MC}} = \begin{cases} \text{Master-MC } V^{F1}, & \text{falls } ((i = k) \wedge (j = l)) \\ 0, & \text{sonst} \end{cases} \quad (\text{Gl. 3.92})$$

mit $\varphi' \in \{0^\circ, 60^\circ, 120^\circ, 180^\circ, 240^\circ, 300^\circ\}$

Feeding-Gewicht Glättung ${}^{\text{M-Smooth}}V_0^{F1}, {}^{\text{M-Smooth}}V_1^{F1}, {}^{\text{M-Smooth}}V_2^{F1}$	2; 1; 0,5
Feeding-Zeitkonstante ${}^{\text{M-Smooth}}\tau^{F2}$	10 Bin
Schwellenverstärkung Orientierungsdetektor ${}^{\text{M-Smooth}}V^\Theta$	8
Schwellenzeitkonstante ${}^{\text{M-Smooth}}\tau^\Theta$	3 Bin
Schwellenoffset ${}^{\text{M-Smooth}}\Theta_0$	4
Feeding-Gewicht Bewegungskontrastdetektor I ${}^{\text{MC-I}}V^{F1}$	3
Inhibition-Gewicht Bewegungskontrastdetektor I ${}^{\text{MC-I}}V^I$	0,5
Feeding-Gewicht Bewegungskontrastdetektor II ${}^{\text{MC-II}}V^{F1}$	0,5
Linking-Gewicht Bewegungskontrastdetektor II ${}^{\text{MC-II}}V^L$	0,5
Feeding-Zeitkonstante ${}^{\text{MC}}\tau^{F1}$	10 Bin
Inhibition-Zeitkonstante ${}^{\text{MC-I}}\tau^I$	10 Bin
Linking-Zeitkonstante ${}^{\text{MC-II}}\tau^L$	10 Bin
Schwellenverstärkung Bewegungskontrastdetektor ${}^{\text{MC}}V^\Theta$	8
Schwellenzeitkonstante ${}^{\text{MC}}\tau^\Theta$	3 Bin
Schwellenoffset ${}^{\text{MC}}\Theta_0$	4
Feeding-Gewicht Master-Bewegungskontrast ${}^{\text{Master-MC}}V^{F1}$	1,5
Feeding-Zeitkonstante ${}^{\text{Master-MC}}\tau^{F1}$	10 Bin
Schwellenverstärkung Master-Bewegungskontrast ${}^{\text{Master-MC}}V^\Theta$	8
Schwellenzeitkonstante ${}^{\text{Master-MC}}\tau^\Theta$	3 Bin
Schwellenoffset ${}^{\text{Master-MC}}\Theta_0$	4

Tab. 3.9: Parameter der Bewegungskontrastdetektoren

3.5.2 Simulationen zur Bewegungskontrastextraktion

Der in Abbildung 3.110 dargestellte Stimulus besteht aus zwei rechteckigen Gebieten, in denen sich räumlich gleichmäßig verteilte Flecken ($g_1=150$, $g_2=50$) mit einer Geschwindigkeit von $v = 4$ Pixel/Bild bewegen. Das kleinere, sich in der Mitte des Stimulus befindende Gebiet verdeckt das größere. Die Flecken bewegen sich im größeren Gebiet nach rechts, währenddessen sich die Flecken des kleinen Gebietes nach links bewegen. Menschliche Beobachter erkennen hierbei die rechteckige Form der beiden Gebiete. Eine Objektdefinition kann in diesem Beispiel nicht durch Intensitätskontrast erfolgen, da beide Bereiche den gleichen mittleren Grauwert besitzen. Die Objektseparation kann aber, wie in Abbildung 3.111 dargestellt, aufgrund der unterschiedlichen Bewegungsrichtungen erfolgen. Das innere Objekt wird in der 180° -Richtungsdetektorschicht repräsentiert, wohingegen die Repräsentation des äußeren Objektes durch die 0° -Richtungsdetektoren erfolgt. Es entstehen somit zwei Bewegungsfelder. Die Form der beiden rechteckigen Gebiete kann durch die Bewegungskontrastdetektion ermittelt werden. In dieser Simulation wurde dazu der Bewegungskontrast I verwendet. Das Ergebnis dieser Bewegungskontrastextraktion ist in Abbildung 3.112 dargestellt.

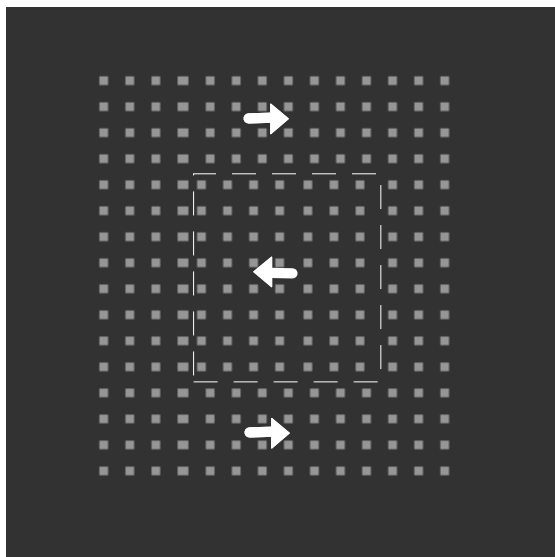


Abb. 3.110: Stimulus zur Bewegungskontrastdetektion (Rechtecke). Die Abbildung zeigt zwei mit Flecken (4 Pixel x 4 Pixel groß) bedeckte Gebiete. Das kleinere Gebiet, das in der Abbildung zur besseren Veranschaulichung durch gestrichelte Linien begrenzt wird, überdeckt das größere Gebiet. Im kleineren Gebiet (ortsfest) bewegen sich die Flecken nach links, im größeren nach rechts. Menschliche Beobachter können hierbei die rechteckigen Formen der beiden Gebiete erkennen.

Die kumulierte Aktivität der Modellneuronen der Master-Bewegungskontrastschicht zeigt die Umrandungen zweier Rechtecke. Diese resultieren aus der räumlichen Bandpaßfilterung der geglätteten Richtungsdetektoraktivitäten. Innerhalb eines Bewegungsfeldes führt die Zentrum/Umfeld-Verschaltungsstruktur zu einem unterschwelligen Membranpotential. Am Rand eines solchen Bewegungsfeldes ist die Hemmung über das Umfeld schwächer, wodurch die Bewegungskontrastdetektoren aktiv werden. Das äußere Rechteck hätte auch durch Intensitätskontrastdetektoren der groben Auflösung detektiert werden können. Die Ränder des kleinen Rechtecks hingegen nicht. Die horizontalen Ränder der Bewegungsfelder werden besser detektiert als die vertikalen, da durch die hexagonale Unterabtastung die räumliche Glättung der Richtungsdetektoraktivitäten zu einem geringeren Gradienten führt, der wiederum eine schwächere Bewegungskontrastdetektoraktivität mit sich bringt.

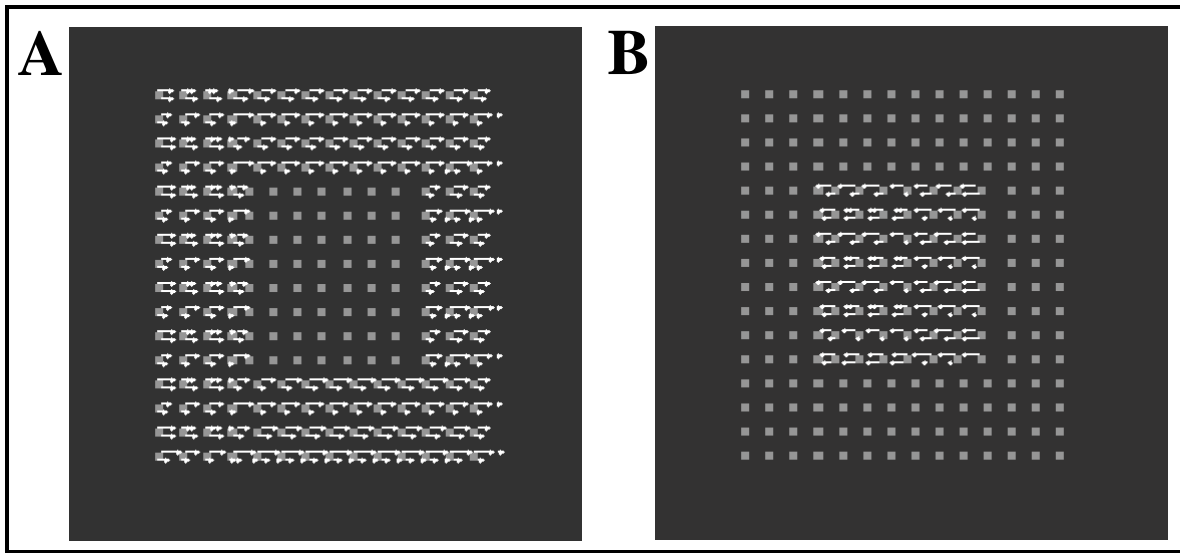


Abb. 3.111: Repräsentation der Vierecke in den Richtungsdetektorschichten. Typische Aktivitäten der 0°- (A) und 180°- (B) ON-Richtungsdetektoren der mittleren Auflösung ($d = 4$ Pixel).

Die rechte Seite des kleinen Rechtecks wird räumlich verteilter repräsentiert als die linke, da die Flecken auf dieser Seite auseinanderlaufen und somit eine kleine Lücke zwischen den beiden Bewegungsfeldern entsteht. Diese Lücke führt dann zu einer Verschmierung des Bewegungskontrastes. Die rezeptiven Felder der aktiven 0°- und 180°-Bewegungskontrastdetektoren liegen auf dieser Seite weiter auseinander als auf der linken.

Die Rechtecke können aufgrund der unterschiedlichen Bewegungsrichtung ihrer Flecken voneinander separiert werden. In Kapitel 2.1.3 wurde aber bereits ein Beispiel mit gegenläufig rotierenden Kreisscheiben gezeigt, bei denen nur aufgrund der absoluten Bewegungsrichtung keine Separation möglich ist. In den nächsten Simulationen wird ein solcher Stimulus verwendet. Die Abbildung 3.113 zeigt zwei mit Flecken ($g_1=150$, $g_2=50$) bedeckte Kreisscheiben, die gegenläufig rotieren. Die Aufgabe des neuronalen Netzwerkes ist die Separation der beiden Kreisscheiben.

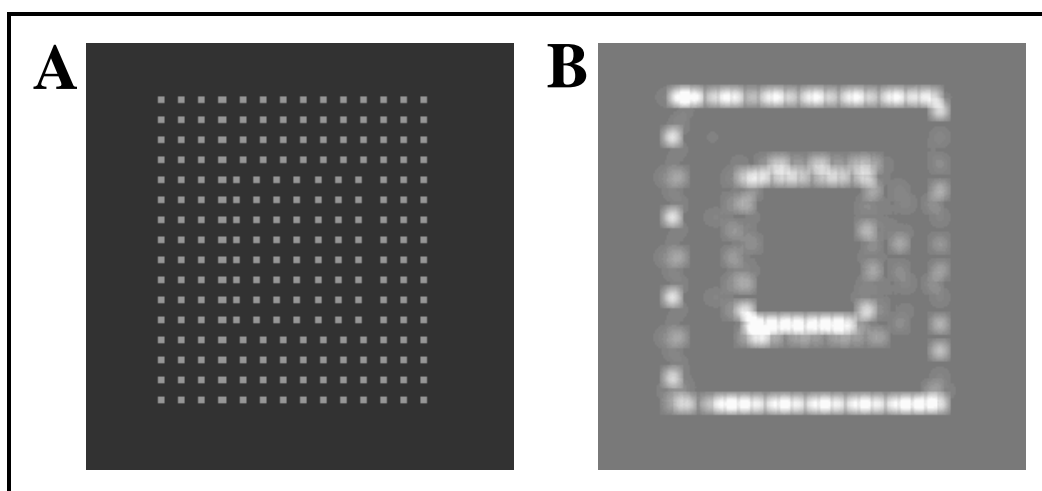


Abb. 3.112: Bewegungskontrastdetektion I („Rechtecke“). Die Abbildung zeigt in A den verwendeten Stimulus aus Abbildung 3.110 und in B den resultierenden Bewegungskontrast. Für die Visualisierung der Modellneuronenaktivitäten der Master-Bewegungskontrastschicht (27×27 Abtastpunkte) wurden Gaußmasken der groben Auflösung ($d = 8$ Pixel) in den Bildbereich zurückprojiziert (Kopecz et al. 1996).

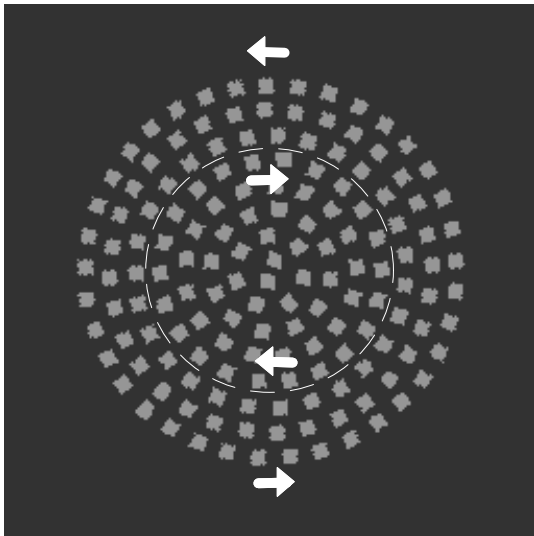


Abb. 3.113: Stimulus zur Bewegungskontrastdetektion („Kreise“). Die Abbildung zeigt zwei gegenläufig rotierende Kreisscheiben (durch graue Rechtecke symbolisiert und durch gestrichelte Linien abgegrenzt). Die innere Scheibe rotiert im Uhrzeigersinn, die äußere gegen den Uhrzeigersinn. Menschliche Beobachter sind in der Lage, die beiden Kreisscheiben zu separieren. Dies kann nicht aufgrund der absoluten Bewegungsrichtung der Flecken erfolgen, da beide Kreisscheiben alle Bewegungsrichtungen beinhalten.

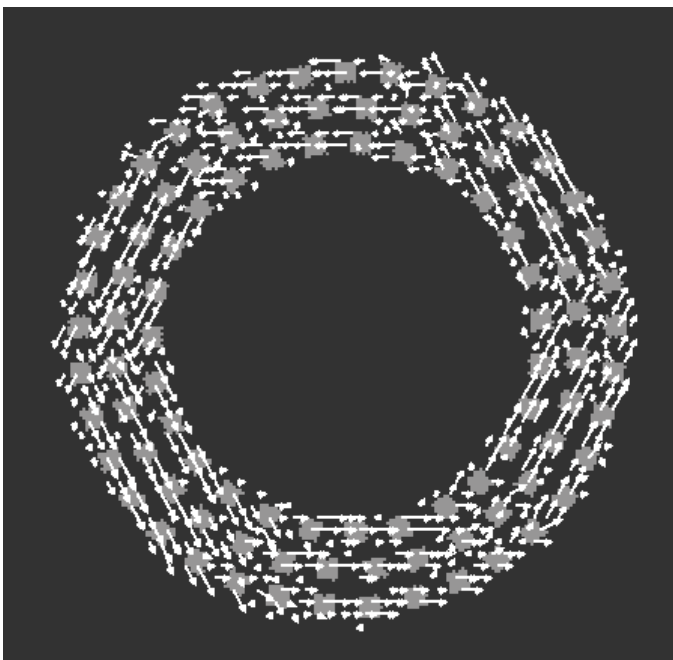


Abb. 3.114:
Repräsentation der äußeren Ringe in den Richtungsdetektorschichten. Typische Aktivitäten der ON-Richtungsdetektoren der mittleren Auflösung ($d = 4$ Pixel). Die neuronale Repräsentation ist auf alle Richtungsdetektorschichten verteilt, wodurch keine Gruppierung über die absolute Bewegungsrichtung möglich ist.

Die Abbildung 3.114 zeigt die Aktivität der ON-Richtungsdetektoren bei Stimulation mit der größeren Kreisscheibe. An diesem Beispiel ist deutlich zu erkennen, daß aufgrund der absoluten Bewegungsrichtung die Flecken nicht zu einer Kreisscheibe gruppiert werden können. Eine genauere Betrachtung der Aktivität zeigt allerdings nur geringe - das sind aufgrund des minimalen Richtungspräferenzunterschiedes in diesem Modell 60° - lokale Bewegungsrichtungsänderungen. Eine Bewegungskontrastdetektion mit der in den Gleichungen 3.88 und 3.89 beschriebenen Verschaltung extrahiert, wie das in Abbildung 3.115 B dargestellte Simulationsergebnis zeigt, Innen- und Außenrand des Objektes. Der mittlere Ring des Objektes führt zu keiner Aktivierung der Bewegungskontrastdetektoren, da sowohl die Bewegung des inneren als auch die des äußeren Ringes hemmend auf die Bewegungskontrastdetektoren wirken. Ohne die Ausdehnung des inhibitorischen Umfeldes auf die benachbarten Richtungspräferenzen würde es hier zu zusätzlichen Grenzen kommen.

Eine Abgrenzung zwischen den rotierenden Kreisscheiben läßt sich auch mit der zweiten neuronalen Verschaltung, beschrieben in den Gleichungen 3.90 und 3.91, zur Bestimmung des Bewegungskontrastes erreichen. Das Simulationsergebnis ist in Abbildung 3.115 D dargestellt. Diese neuronale Verschaltung führt nur zu einer Trennung der beiden Objekte. Der Rand der größeren Kreisscheibe wird von den Bewegungskontrastdetektoren nicht markiert.

Zur Formerkennung müßte dieser zusätzlich noch von den retinalen Ganglienzellen zugeführt werden.

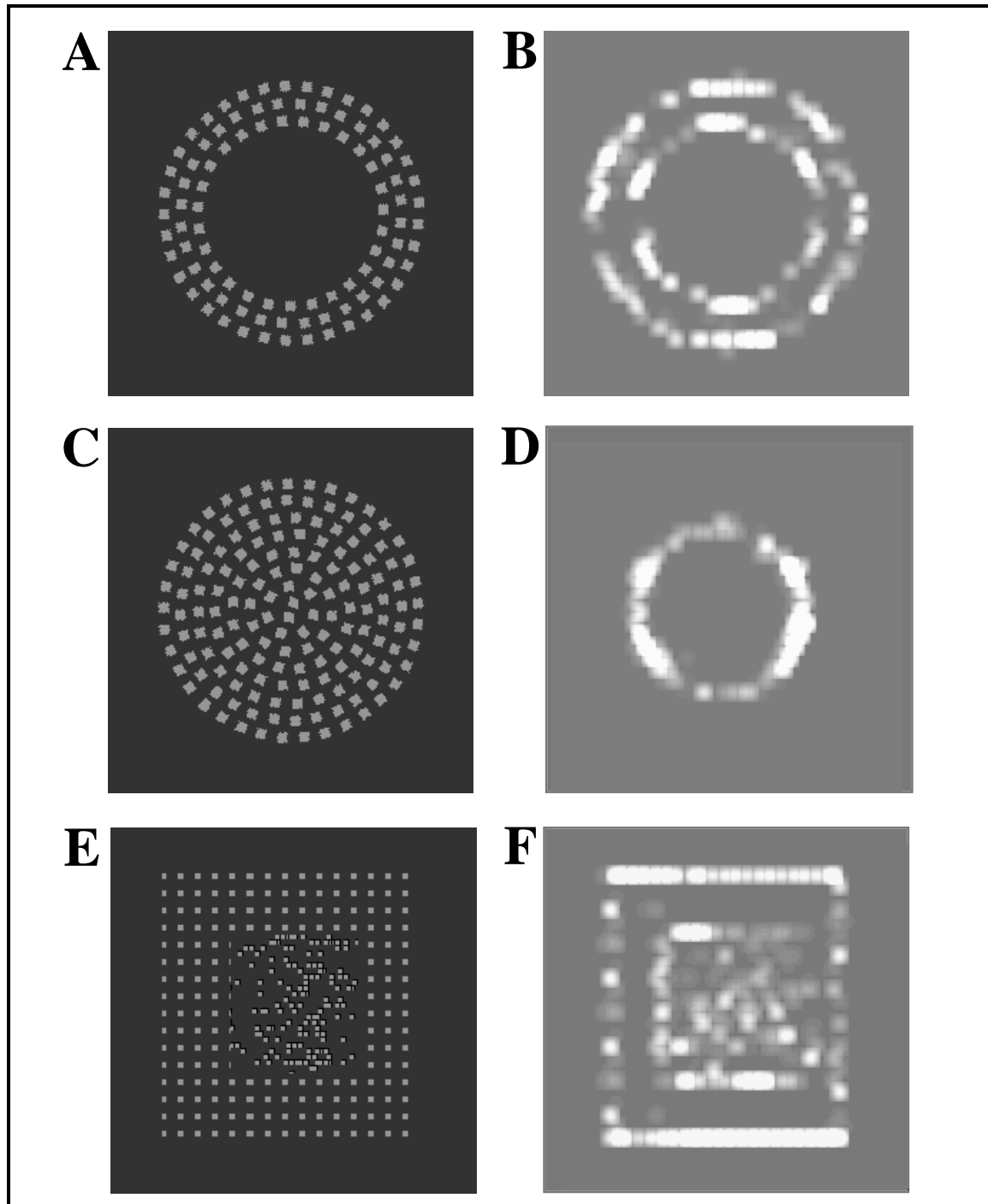


Abb. 3.115: Bewegungskontrastdetektion. Die Abbildung zeigt in der linken Spalte jeweils den verwendeten Stimulus und in der rechten Spalte den resultierenden Bewegungskontrast. **A:** Drei sich gegen den Uhrzeigersinn mit derselben Winkelgeschwindigkeit drehende Ringe. **B:** Markierung der inneren und äußeren Objektgrenze durch Bewegungskontrast I. **C:** Stimulus A mit zusätzlichen Ringen, die sich mit dem Uhrzeigersinn drehen. **D:** Der resultierende Bewegungskontrast II trennt die beiden Objekte voneinander. Der äußere Rand der größeren Kreisscheibe wird nicht detektiert. **E:** Stimulus wie in Abbildung 3.110, aber mit sich in zufälliger Richtung bewegendes Flecken im inneren Rechteck. **F:** Der Bewegungskontrast I kann die Form des inneren Rechtecks nicht mehr gewinnen. Gleiches gilt natürlich für den Bewegungskontrast II. Zur Bestimmung der Objektkanten sind hier weitere Verarbeitungsstufen, wie zum Beispiel der Kontrast des Bewegungskontrastes, nötig.

Ein Vergleich der Verschaltungen zeigt bei beiden Vor- und Nachteile. Die erste Verschaltung führt zur Gewinnung der Form des Objektes auch bei Fehlen anderer bewegter Objekte in dessen räumlicher Umgebung. In der Abbildung 3.116 ist für diesen Fall eine Bewegungskontrastdetektion anhand eines Beispiels mit einer realen Szene dargestellt. Die Form der sich bewegenden Fußgängerin wird durch den Bewegungskontrast extrahiert. Ein Nachteil dieser Verschaltung ist die Möglichkeit, daß in Richtungsdetektorschichten unterschiedlicher Auflösung aufgeteilte Bewegungsfelder zu falschen Abgrenzungen führen. Abhilfe schafft dabei nur die Einbeziehung mehrerer Auflösungen in die Bewegungskontrastverarbeitung. Ein Nachteil beider Verschaltungen ist die nicht mögliche Formextraktion des in Abbildung 3.115 E gezeigten Stimulus, bei dem sich die Flecken des inneren Rechtecks nicht mehr homogen nach links, sondern in zufällige Richtungen bewegen. In diesem Beispiel müßte sich zusätzlich noch eine Kontrastdetektion an die Bewegungskontrastdetektion anschließen.

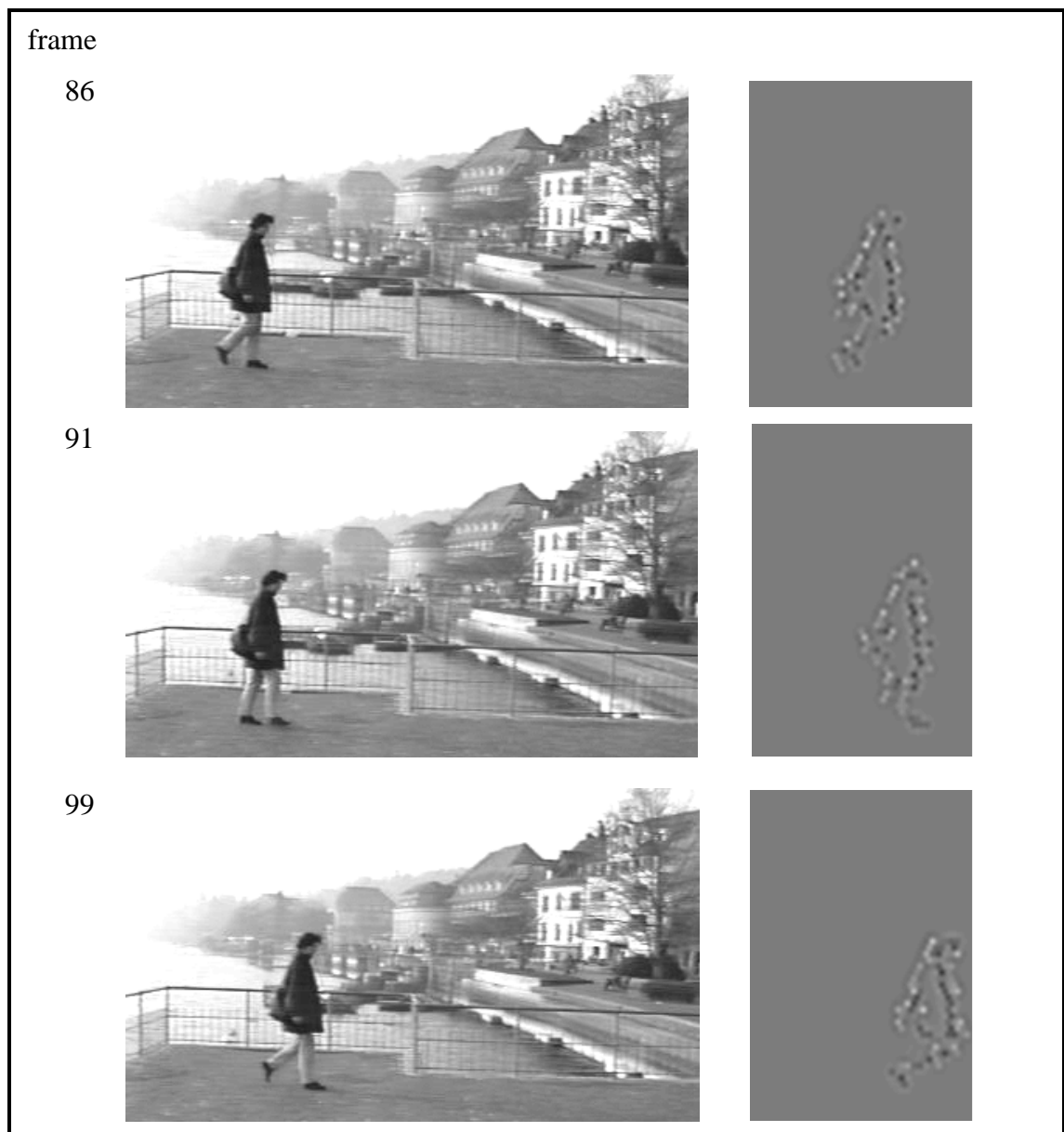


Abb. 3.116: Bewegungskontrastdetektion I in einer realen Szene. Eine Fußgängerin bewegt sich vor ruhendem Hintergrund nach rechts. Das neuronale Netzwerk extrahiert die Form der Fußgängerin durch Bestimmung des Bewegungskontrastes I. Zu beachten ist, daß das Standbein nicht von den Bewegungskontrastdetektoren detektiert wird. Für eine vollständige Gruppierung ist eine Kopplung von Kontur/Form-System und Transientensystem notwendig. (Abbildung nach Schott et al. 1996.)

4 Diskussion

Das in dieser Arbeit entwickelte neuronale Netzwerk basiert auf einem allgemeinen Ansatz und kann flexibel eingesetzt werden. Die neuronalen Verschaltungen werden durch neurophysiologische und psychophysische Untersuchungen des Sehsystems motiviert. Das neuronale Netz kann unter Verwendung derselben Parameter sowohl Wahrnehmungsaspekte erklären als auch zur technischen Bildverarbeitung eingesetzt werden. Zudem kann es zur Beschreibung der Informationsverarbeitung im biologischen Sehsystem verwendet werden. In diesem Kapitel sollen die Leistungen und die Struktur des entwickelten neuronalen Netzes zur Bewegungsdetektion mit Ergebnissen psychophysischer und neurophysiologischer Untersuchungen des Sehsystems verglichen werden. Außerdem soll auf die Relevanz des entwickelten neuronalen Netzes für die technische Bildverarbeitung eingegangen werden.

Bei der Simulation der psychophysischen Effekte wird davon ausgegangen, daß die menschliche Bewegungswahrnehmung sich in der raum-zeitlichen Aktivitätsverteilung der Neuronen widerspiegelt (z.B. Salzman et al. 1990, Nakamura et al. 1994) und somit die Aktivierung der Modellneuronen als Kriterium für die Bewertung der Simulationen herangezogen werden kann.

4.1 Vorverarbeitung

Die Funktionsweise der Vorverarbeitung entspricht in einigen Eigenschaften der der Retina. In Analogie zum Sehsystem von Primaten oder von Katzen wird im vorgestellten Modell im Zentrum eine hochauflösende und in der Peripherie eine grobauflösende Verarbeitung durchgeführt. Im Gegensatz zur biologischen Retina nehmen die RF-Größen nicht kontinuierlich mit der Exzentrizität zu, sondern es gibt drei unterschiedlich große, sich überlappende Verarbeitungsbereiche (Kap. 3.1.1.2). Die grobe Auflösung soll hierbei lediglich Veränderungen in der Peripherie des Gesichtsfeldes extrahieren, damit die Kamera auf diese schwenken und ein Kontur/Form-System mit dem hochauflösenden Verarbeitungsbereich diese analysieren kann. Die mittlere Auflösung kann gleichermaßen für das Transientensystem und das Kontur/Form-System verwendet werden.

Die Verarbeitung der Eingabebilder erfolgt durch Rezeptoren, Bipolarzellen und Ganglienzellen. Aus den Antworten der Rezeptoren werden Intensitätskontraste durch Bipolarzellen des ON- und OFF-Pfades gewonnen. Anschließend werden die graduierten Potentiale durch Modellneuronen, die den retinalen Ganglienzellen entsprechen, in Folgen von Aktionspotentialen umgesetzt. Dabei kann je nach Wahl der Parameter ein X-Zellen-Verhalten oder ein Y-Zellen-Verhalten modelliert werden (Kap. 3.1.1.5). Da für ein bewegungssensitives System besonders zeitliche Änderungen wichtig sind, wurden die Parameter der Modellneuronen so gewählt, daß das Antwortverhalten einen starken phasischen (zeitlicher Bandpaß) und nur einen sehr geringen tonischen (zeitlicher Tiefpaß) Anteil aufweist.

In der Vorverarbeitung wurde großer Wert auf eine geringe Anzahl an generierten Aktionspotentialen gelegt, denn die Übertragungsrate von der Vorverarbeitung auf das Accelerator-System soll unter Berücksichtigung einer schnellen Verarbeitung gering sein. Die Gesamtaktivität der ON- und OFF-Ganglienzellen der feinen Auflösung liegt zum Beispiel bei der Sequenz „Durlacher Tor“ (Kap. 3.1.2.2) deutlich unter einem Prozent und genügt damit dieser Anforderung.

Problematisch erweist sich die Detektion von Kanten mit geringem Intensitätskontrast. Im Hinblick auf die nachgeschalteten Richtungsdetektoren sollte sich die Anzahl an generierten Aktionspotentialen bei niedrigen und hohen Intensitätskontrasten nicht zu sehr unterscheiden und nicht zu gering sein. Um Kanten mit niedrigem Intensitätskontrast in ausreichend großer Anzahl an Aktionspotentialen zu kodieren, sind mehrere Verfahren denkbar:

Benachbarte Ganglienzellen, die eine Kante detektieren, könnten sich über laterale Verbindungen gegenseitig verstärken. Da die laterale Verschaltung isotrop sein muß und die Verstärkung nur auf Ganglienzellen mit rezeptiven Feldern, die entlang der Kante angeordnet sind, beschränkt sein soll, bietet sich hier die Verwendung von Linking-Verbindungen an (Becker 1995). Ein technischer Nachteil dieses Verfahrens ist die aufgrund der Isotropie rasch ansteigende Verbindungsanzahl mit zunehmender Reichweite und die zur Aktivitätsbeschränkung zusätzlich zu implementierende Inhibition. Aus Gründen der Rechenzeit und der Möglichkeit, das neuronale Netzwerk der Vorverarbeitung parallelisieren zu können, wurde auf laterale Interaktionen verzichtet.

Rückwärts gerichtete Linking-Verbindungen von Orientierungsdetektoren mit orientierten RF auf Zellen mit konzentrischen RF könnten die Detektion von Kanten mit schwachem Intensitätskontrast unterstützen (Behring 1998). Da die Ganglienzellen sich in der Vorverarbeitung, die Orientierungsdetektoren sich hingegen auf dem Accelerator-System befinden, läßt sich eine solche Rückkopplung praktisch nicht realisieren. Ein Duplikat der Ganglienzellschicht auf dem Accelerator würde dieses Problem lösen, doch dies würde zusätzliche Modellneuronen auf dem Accelerator binden. Weiterhin entstünde ein technisches Übertragungsproblem von der Vorverarbeitung zum Accelerator, da die Intensitätskontraste in der Aktionspotentialdichte kodiert werden müßten, welche bei hohen Intensitätskontrasten zu groß für eine Übertragung unter Echtzeitanforderungen wäre.

Ein zusätzlicher Adaptationsmechanismus im Modellneuron durch eine zweite dynamische Schwelle mit großer Schwellenzeitkonstante könnte auch zur Aktivitätsbeschränkung herangezogen werden (z.B. Lankheet et al. 1989, Czambor 1997). Eine zweite dynamische Schwelle würde aber die Rechenzeit erheblich verlängern. Das gleiche gilt für einen Adaptationsmechanismus, der die Aktivität der Ganglienzellen mit benachbarten rezeptiven Feldern miteinbezieht.

Bei der Verwendung einer schnellen Hardware sollte in einer Weiterentwicklung des Vorverarbeitungsmoduls eine oder mehrere der genannten Möglichkeiten implementiert werden.

4.2 Detektion der Konturbewegungsrichtung

4.2.1 Einordnung in psychophysische Untersuchungen

Im Vordergrund dieser Arbeit stand die Frage, wie - unter Berücksichtigung des Apertur- und Korrespondenzproblems - das menschliche Sehsystem zu einer eindeutigen Bewegungswahrnehmung und zu einer Objektseparation aufgrund von Bewegungsinformation gelangen könnte. Aus den Ergebnissen psychophysischer Untersuchungen konnten Vorschläge für mögliche neuronale Verschaltungen gewonnen werden. Das aus diesen Überlegungen entstan-

dene neuronale Netz zeigt mögliche Verschaltungsstrukturen, mit denen zahlreiche auf Bewegung basierende psychophysische Effekte erklärt werden können.

Ein einzelner bilokaler Richtungsdetektor unterliegt dem Apertur- und Korrespondenzproblem. Die vom Richtungsdetektor detektierte Bewegungsrichtung kann von der physikalischen um bis zu $\pm 90^\circ$ abweichen, was zur Folge hat, daß mehrere Richtungsdetektoren mit gleichen RF-Mittelpunkten, aber unterschiedlicher Richtungspräferenz aktiv sind.

Die psychophysischen Ergebnisse von z.B. Chang und Julesz (1984) sowie Yuille und Grzywacz (1988) zeigen eine räumliche (*motion capture*, siehe Kap. 2.1.2.3) und die Untersuchungen von z.B. Anstis und Ramachandran (1987) und Pinkus und Pantle (1997) eine zeitliche (*priming*, siehe Kap. 2.1.2.2) Integration der Bewegungsinformation. In der neuronalen Verschaltung finden sich diese in den lateralen Linking-Verbindungen zwischen den benachbarten Richtungsdetektoren wieder. In Kapitel 3.2.2.1 konnte gezeigt werden, daß mit Hilfe unverzögerter und verzögerter Linking-Kopplungen und lokaler Inhibition über Interneuronen diese psychophysischen Effekte erklärt werden können und das Ergebnis der Richtungsdetektion eindeutig wird. Mittels Voraktivierung ist es zum Beispiel möglich, einen einzelnen geradlinig bewegten Flecken in einem sich in den Richtungen zufällig, aber nicht geradlinig bewegenden Fleckenmuster zu detektieren, da die zeitliche Pfad-Integration der Bewegungsinformation zu einer hohen Aktivität der entsprechenden Richtungsdetektoren führt (Abb. 3.38). Die zeitliche und räumliche Integration der Bewegungsinformation führt aber nicht nur zu einer eindeutigen Bewegungsrichtungsdetektion, sondern sie erhöht zugleich die Robustheit des Systems gegen Rauschen.

Das Modell berücksichtigt nicht nur die Notwendigkeit einer räumlichen und zeitlichen Integration der Bewegungsinformation, es besitzt auch getrennte ON- und OFF-Verarbeitungspfade. Die Trennung in ON- und OFF-Pfad ermöglicht die Unterstützung der Korrespondenz aufgrund der Intensitäts-Polarität (Shechter und Hochstein 1990, siehe Kap. 2.1.2.1). Das hier vorgestellte neuronale Netz zur Richtungsdetektion kann allerdings den absoluten Grauwert nicht nutzen. Der absolute Grauwert trägt nach Shechter und Hochstein (1990) auch nur zur Unterstützung der Korrespondenz bei, wenn die Grauwerte sehr unterschiedlich sind.

4.2.2 Einordnung in biologische Untersuchungen

Die neurophysiologischen Untersuchungen von Albright (1984) zeigten einen großen Anteil (84 %) von MT-Neuronen mit und nur einen kleinen Anteil (8 %) ohne Richtungsselektivität. In dem Modell zur Bewegungsdetektion sind sechs von sieben Modellneuronen richtungsselektiv. Die inhibitorisch wirkenden Interneuronen besitzen keine Richtungsselektivität. Das Verhältnis von Modellneuronen mit und ohne Richtungsselektivität entspricht somit neurophysiologischen Ergebnissen.

Die hier vorgestellten Richtungsdetektoren besitzen zur Bestimmung der Richtung zwei Eingänge, von denen einer verzögert wird. Solche Verzögerungen kommen auch im biologischen Sehsystem vor, wie zum Beispiel die Untersuchungen von Dinse und Krüger (1994) zeigten. Bereits Zellen des CGL (Katze) antworten nach dem Anschalten des Stimulus mit Unterschieden in den Verzögerungen von bis zu 30 ms. Dies entspricht ungefähr der in diesem Modell implementierten Verzögerungen. Die Richtungsselektivität eines biologischen Neurons könnte somit durch eine Verschaltung von CGL-Zellen mit unterschiedlichen Verzögerungen in ihrem zeitlichen Antwortverhalten, wie in diesem Modell implementiert, erfolgen.

Die Ergebnisse von Movshon et al. (1978) und Mikami et al. (1986) führten zu der Annahme, daß fazilitatorisch und inhibitorisch wirkende Mechanismen zur Richtungsselektivität führen (Kap. 2.2.3). Die in diesem Modell verwendeten Richtungsdetektoren sind sowohl fazilitatorisch über Linking-Verbindungen als auch inhibitorisch über Interneuronen gekoppelt. Die Richtungsselektivität beruht allerdings nicht nur auf diesen beiden Mechanismen, sondern geht in erster Linie auf die zeitlich unterschiedliche Kopplung von Zellen mit räumlich benachbarten rezeptiven Feldern zurück. Die lateralen Kopplungen über Linking-Verbindungen tragen in Verbindung mit einer inhibitorischen Rückkopplung über Interneuronen erheblich zu einem eindeutigen Ergebnis der Bewegungsrichtungsbestimmung bei. Mikami entdeckte 1992 mit Folgen von Lichtblitzen als Reiz, die so angeordnet waren, daß ein geradliniger Bewegungseindruck entstand, eine Abhängigkeit der Aktivität der MT-Neuronen von der Anzahl der Lichtblitze. Je mehr Lichtblitze dargeboten wurden, um so stärker waren die MT-Neuronen aktiv. Dieses Ergebnis ist konform mit den psychophysischen Untersuchungen zur zeitlichen Integration der Bewegungsinformation. Die in dieser Arbeit verwendeten Modellneuronen mit Richtungspräferenz besitzen ebenfalls diese Eigenschaft. Durch die Voraktivierung steigt bei geradliniger Bewegung die Aktivität bis zu einem Maximalwert, wie die Simulation mit dem geradlinig bewegten Flecken in dem zufällig bewegten Fleckenmuster zeigte (Abb. 3.38).

Die Richtungspräferenz der Modellneuronen ähnelt denen der von Maunsell und van Essen 1983 untersuchten MT-Neuronen. Die Richtungscharakteristik der Modellneuronen hat eine Halbwertsbreite von 60° , die der MT-Neuronen ist etwas größer. Beide Richtungscharakteristiken sind ungefähr gleich scharf.

4.2.3 Vergleich mit anderen Modellen

Der in dieser Arbeit entwickelte Richtungsdetektor beinhaltet Eigenschaften mehrerer verschiedener Modelle. Er basiert auf dem von Reichardt (1957) vorgestellten Bewegungsdetektor, lehnt sich aber auch an die zeitliche und räumliche Integrationsfähigkeit anderer Modelle an (Barlow und Levick 1965, Glünder und Lehmann 1992, siehe Kap. 2.3.1). Ein Marburger Modellneuron wird zum bilokalen Richtungsdetektor, indem es von zwei retinalen Ganglienzellen des Y-Pfades verzögert und unverzögert Eingabe erhält. Im Gegensatz zu dem Reichardt-Detektor werden die zeitlichen Intensitätsänderungen vor der Bewegungsdetektion detektiert, was zur Verringerung des Korrespondenzproblems beiträgt. Reichardt vermeidet hingegen die Detektion stehender Objekte durch die Verrechnung entgegengesetzter Bewegungsrichtungen.

Der hier vorgestellte Richtungsdetektor hat im Gegensatz zu richtungssensitiven Orientierungsdetektoren zwei kleine konzentrische Teilfelder. Orientierungsdetektoren mit asymmetrischen rezeptiven Feldern können zwar auch Bewegungsrichtungen von Kanten, nicht aber von einzelnen Flecken detektieren (Kap. 3.4.1.1). Da in dieser Arbeit auch psychophysische Effekte mit Flecken simuliert werden sollten, wurde ein bilokaler Richtungsdetektor entwickelt. Ein Vorteil bei der Verwendung von Orientierungsdetektoren als Richtungsdetektoren wäre das verringerte Korrespondenzproblem, da die Kanten vor der Richtungsdetektion in getrennte neuronale Schichten mit unterschiedlicher Orientierungspräferenz aufgeteilt würden. An jede Orientierungsdetektorschicht müßten sich aber jeweils sechs Richtungsdetektorschichten und eine Schicht Interneuronen anschließen, was zu einer extrem hohen Anzahl an Modellneuronen führen würde.

Die Voraktivierung benachbarter Richtungsdetektoren lehnt sich an den Bewegungsdetektor von Barlow und Levick (1965) an. Im Vergleich zu deren Detektor besteht bei dem vorgestellten Richtungsdetektor bereits eine Richtungsselektivität ohne laterale Kopplung.

Durch die Bilokalität der Richtungsdetektoren wird mit Hinblick auf eine eindeutige lokale Detektion bereits die Anzahl an aktiven Modellneuronen reduziert. Die Voraktivierung dient in diesem neuronalen Netz insbesondere der Reduzierung des Korrespondenzproblems. Dies gilt in gleichem Maße für die unverzögerten, lateralen Verschaltungen. Der von Glünder und Lehmann (1992) beschriebene mathematische Relaxationsprozeß wird in diesem Modell durch einen dynamischen, biologisch plausiblen Interaktionsprozeß zwischen impulskodierenden Modellneuronen ersetzt.

An den Richtungsdetektoren soll ohne eine weitere Analyse direkt die Bewegungsrichtung bestimmt werden können. Deswegen wurde auf ein eindeutiges Ergebnis bei der Richtungsdetektion Wert gelegt. Dies ermöglicht zum Beispiel eine einfache Verschaltungsstruktur bei der Detektion des Bewegungskontrastes oder der Verschiebung des Aufmerksamkeitsfokus (Pauly et al. 1997, 1998). Nachteilig ist hierbei zweifellos die Diskretisierung der Bewegungsrichtung in 60° Abständen, so daß eine genaue Bewegungsrichtung nicht ermittelt werden kann.

Ferner ist zu beachten, daß das Transientensystem alleine nicht immer die vollständige Kontur eines Objektes extrahieren kann (z.B. Abb. 3.24). Das Detektionsergebnis hängt maßgeblich von der Ausrichtung der Objektkanten zur Bewegungsrichtung ab. Parallel zur Bewegungsrichtung orientierte Kanten werden vom Transientensystem - mit Ausnahme der Linienenden - nicht detektiert. Die gesamte Objektkontur läßt sich im allgemeinen nur durch das Kontur/Form-System oder aus einer Kombination beider Systeme gewinnen.

4.2.4 Motion Capture durch Richtungsdetektoren unterschiedlicher räumlicher Auflösung

Die Herstellung der Korrespondenz wird auch durch unterschiedliche räumliche Auflösungen inklusive Geschwindigkeitsbereiche ermöglicht (Green 1986). Das Modell unterstützt dies durch die Verwendung von drei verschiedenen Auflösungen. Die hochauflösenden Richtungsdetektoren extrahieren bevorzugt kleine Geschwindigkeiten, grobauflösende Richtungsdetektoren große Geschwindigkeiten.

Richtungsdetektoren mit gleichen RF-Mittelpunkten, aber unterschiedlicher Auflösung können mehrdeutig auf eine Kantenbewegung antworten. Die in Kapitel 3.3 dargestellte Kopplung der Auflösung führt zu einem eindeutigen Ergebnis der Bewegungsdetektion. Die Richtungsdetektoren unterschiedlicher Auflösung haben verschiedene Geschwindigkeitscharakteristiken (Abb. 3.40). Stehen die Richtungsdetektoren über eine Kopplung im Wettbewerb, so setzt sich der am stärksten aktivierte Richtungsdetektor durch und das Ergebnis der Bewegungsdetektion wird eindeutig.

Die Kopplung zwischen Richtungsdetektoren unterschiedlicher Auflösung ermöglicht auch die Erklärung des von Ramachandran und Cavanagh (1987) beobachteten Effekts des *motion captures* zwischen einem sinusförmigen Gitter niedriger Frequenz, das sich mit konstanter Geschwindigkeit bewegt, und kleinen, sich in die entgegengesetzte Richtung bewegenden Flecken. Bei ausreichend großer Geschwindigkeit des sinusförmigen Gitters wird eine der dem Gitter entsprechende Bewegungsrichtung an den kleinen Flecken detektiert. Die Simulationsergebnisse unterstützen die Annahme Ramachandrans und Cavanaghs, daß die wahrgenommene Bewegungsrichtung durch die Bewegung der niederfrequenten Objektanteile dominiert wird.

4.2.5 Objektverfolgung

Das vorgestellte neuronale Netz kann auch zur technischen Bildverarbeitung eingesetzt werden. Ein Transientensystem hat in einem technischen Gesamtsystem einerseits die Aufgabe, eine Szenensegmentierung mittels Bewegungskontrastdetektion durchzuführen, andererseits soll es Bildbereiche markieren, auf die ein hochauflösender, aber kleiner Verarbeitungsbereich des Kontur/Form-Systems (z.B. Weitzel 1998) gesetzt werden kann. Da die Analyse des Bildbereiches durch ein Kontur/Form-System mehrere hundert Millisekunden dauert, muß der hochauflösende Bearbeitungsbereich bei sich bewegenden Objekten nachgeführt werden, wozu die Bewegungsrichtung des Objektes benötigt wird. Das Versetzen des hochauflösenden Verarbeitungsbereiches, das einer Sakkade entspricht, und das Verfolgen von Objekten mit diesem Verarbeitungsbereich wird von einer Aufmerksamkeitssteuerung durchgeführt (z.B. Pauly et al. (1997), zur „modellbasierten Objektverfolgung“ mit dynamischen neuronalen Netzen siehe Lindemann (1998)).

Das Transientensystem verfügt über hochauflösende Richtungsdetektoren. Obwohl die rezeptiven Felder der Neuronen im magnozellularen Verarbeitungspfad in der Regel größer sind als die des parvozellularen Verarbeitungspfades, gibt es nach Lagae et al. (1993) Neuronen, die besonders selektiv für kleine Geschwindigkeiten sind. Ihre rezeptiven Felder befinden sich in der Fovea, so daß diese Neuronen wahrscheinlich für die Objektverfolgung verantwortlich sind. Die Richtungsdetektoren der hohen Auflösung können kleine Geschwindigkeiten detektieren, so daß ihre Aktivität dazu benutzt werden kann, ein sich bewegendes Objekt zur genaueren Analyse mit dem Aufmerksamkeitsfokus verfolgen zu können (Pauly et al. 1998).

4.3 Barberpole und Barberplaid Illusion

4.3.1 Einordnung in psychophysische Untersuchungen

Einen starken Hinweis auf die richtige Bewegungsrichtung geben Terminatoren (Hildreth 1984, Nakayama und Silverman 1988b, siehe Kap. 2.1.2.4). Deswegen wurde das neuronale Netz zur Bewegungsrichtungsdetektion in Kapitel 3.4 noch um einen zusätzlichen Pfad erweitert, in dem Linienenden und deren Bewegungsrichtung extrahiert werden. Die gewonnene Bewegungsrichtung der Linienenden wird anschließend zur Beeinflussung der Bewegungsrichtungsdetektion an der zugehörigen Kontur verwendet. Damit kann die Mehrdeutigkeit in der Bewegungsdetektion an der Kontur verringert werden. Mit dem vorgestellten Modell zur Einbindung der Linienenden in die Bewegungsdetektion können sowohl die *barberpole illusion* als auch die *barberplaid illusion* erklärt werden. Die Störung der *barberpole illusion* durch die von Kooi 1993 vorgeschlagenen Einschnitte in der Apertur (Abb. 2.9) konnten ebenfalls mit diesem Modell gezeigt werden. Bei den Simulationen wurde auf den Einfluß der Tiefeninformation aus Gründen der Modellneuronenanzahl verzichtet und davon ausgegangen, daß alle extrahierten Linienenden intrinsisch sind, d.h. zu dem Objekt gehörig. Eine Erweiterung des vorliegenden Modells müßte somit die Entscheidung treffen, ob die Terminatoren zu dem Objekt gehören oder ob sie aus einer Verdeckung resultieren. Daran anschließend müßte die Kopplung zwischen Richtungsdetektoren für Linienenden und Richtungsdetektoren für Konturen gesteuert werden. Zur Tiefendetektion könnten bei diesem „monokularen“ System T-Stücke und Balkenreizbreiten herangezogen werden. Eine mögliche Szenensegmentierung aufgrund von Tiefeninformation wurde unter Verwendung von impulskodierenden neuronalen Netzen bereits von Ostkamp (1996) und Weitzel (1998) mit

Hilfe von T-Stücken und von Saam et al. (1997) mittels Intensitätskontrast und Kopfbewegungsparallaxe gezeigt.

Die Wahrnehmung der Bewegungsrichtung wird aber nicht ausschließlich von den intrinsischen Terminatoren dominiert, wie die Untersuchungen von Wilson und Mast (1993) mit Scheinkonturen, die durch Linienenden gebildet werden, zeigen (Abb. 2.10 A). In diesem Zusammenhang ist auch die Verarbeitungsdauer interessant. Lange Stimuluspräsentationszeiten (> 150 ms) rufen eine wahrgenommene Bewegungsrichtung hervor, die aus der Scheinkontur resultiert, kurze (≈ 60 ms) hingegen eine Bewegungsrichtung, die der der Linien und in dieser speziellen Untersuchung der der Linienenden entspricht. Der zeitliche Aktivitätsverlauf (Abb. 3.96) der Richtungsdetektoren bei der Simulation der *barberpole illusion* zeigt, daß die durch die Terminatoren bestimmte Bewegungsrichtung erst nach einer Dauer von etwa 100 ms bis 120 ms dominiert. Wenn die *barberpole illusion* auf der Terminatorbewegungsrichtung beruht, dann sollte die Illusion von der Darbietungsdauer abhängen, da die Extraktion der Linienenden, die Bestimmung ihrer Bewegungsrichtung und die Änderung der neuronalen Dynamik, wie die Simulation gezeigt hat, einer gewissen Verarbeitungszeit bedarf. Eine kurze Präsentation des Stimulus müßte dann zu einer Bewegungswahrnehmung orthogonal zu der Orientierung der Balkenreize führen, wohingegen eine lange Präsentation die Illusion hervorrufen würde.

4.3.2 Einordnung in biologische Untersuchungen

1989 wiesen Rodman und Albright nach ihren Untersuchungen mit „Barberplaid-Stimuli“ auf zwei Typen von MT-Neuronen hin. Die Neuronen des einen Typs waren selektiv für die Bewegung des einzelnen Gitters, die des anderen Typs sensitiv für das gesamte Muster. In dem in dieser Arbeit entwickelten Modell gibt es diese Unterscheidung aus Gründen der Modellneuronenzahl nicht. Nahezu alle Modellneuronen, die ohne die Einbindung der Bewegungsrichtung der Linienenden sensitiv für die einzelne Gitterbewegung waren, wurden unter der Beeinflussung durch Richtungsdetektoren für Linienenden sensitiv für das Gesamtmuster. Daß die Trennung in zwei Typen von MT-Neuronen nicht erwiesen ist, zeigen die Untersuchungen von Stoner und Albright (1992), die einen Wechsel des Antwortverhaltens beobachteten. Dieser Wechsel ist auch in der Simulation festzustellen.

Die Bestimmung der Kantenorientierungen und der Linienenden erfolgt nach dem bereits in den sechziger Jahren von Hubel und Wiesel (1960, 1965, 1968) vorgeschlagenen Verschaltungen. Entsprechend werden Kantenorientierungen durch Modellneuronen mit gerichteten rezeptiven Feldern ermittelt. Diese Modellneuronen sind lateral über Linking-Verbindungen gekoppelt, so daß Modellneuronen, deren rezeptive Felder entlang einer Intensitätskontrastkante liegen, sich gegenseitig unterstützen können, wodurch ein Assoziations-Feld entsteht (Eckhorn et al. 1990). Die Linking-Kopplungen wurden dabei so gewählt, daß sich in ihnen das Gestaltprinzip des guten Verlaufs widerspiegelt. Alle Verbindungen der Orientierungsdetektoren könnten wahrscheinlich, wie auch die der anderen Klassifizierer, mit Hilfe der Hebbschen Lernregel erzeugt werden (Kupper 1998, Thiel 1998, Saam und Eckhorn 1998). Die Verschaltung der Linienendendetektoren entspricht der der Hyperkomplexzellen (Hubel und Wiesel 1960, 1965, 1968), deren gerichtetes rezeptives Feld aus einem exzitatorischen und einem inhibitorischen Teilgebiet besteht, so daß es bei Reizung mit einem das rezeptive Feld völlig überdeckenden Balkenreiz zu keiner Aktivierung des Detektors kommt, wohingegen der Detektor bei einer ausschließlichen Überdeckung des exzitatorischen Teilgebiets aktiv wird.

Neurophysiologisch ist die Richtungsdetektion der Linienenden und ihre Einwirkung auf die Richtungsdetektion der Kontur noch nicht abgesichert. Die Existenz von Linienenden-

detektoren, die zusätzlich eine Richtungssensitivität aufweisen, scheint unter Beachtung der Wahrnehmungsuntersuchungen - wie insbesondere die *barberpole illusion* und *barberplaid illusion* zeigen - wahrscheinlich zu sein. Das vorgestellte neuronale Netz mit impulskodierenden Modellneuronen zeigt, daß die *barberpole* und die *barberplaid illusion* mit der Einbindung der Linienendenbewegung in die Bewegungsverarbeitung erklärt werden können und unterstützt somit diese Vermutung. Die Verschaltung zu solchen richtungssensitiven Linienendendetektoren könnte prinzipiell so erfolgen, wie sie in dieser Arbeit geschildert wurde. Mit dieser neuronalen Verschaltung konnte sogar die Störung der *barberpole illusion* durch Einschnitte in der Apertur, wie sie in den Untersuchungen Koois (1993) vorkommen, gezeigt werden. Allerdings müßten noch zusätzliche Verschaltungen zur Verarbeitung der Tiefeninformation implementiert werden, um die von zum Beispiel Shimojo et al. (1989) sowie Vallortigara und Bressan (1991) gezeigte Störung der *barberpole illusion* durch Tiefenhinweise zeigen zu können.

4.3.3 Vergleich mit anderen Modellen

Im Grundlagenkapitel wurden bereits die Verfahren zur Generierung eines kohärenten Bewegungseindrucks bei der *barberplaid illusion* beschrieben. Nach der Bestimmung der lokalen Gitterbewegungsrichtung kann mittels Vektorsummation, Vektormittelung oder dem IOC-Verfahren eine resultierende Bewegungsrichtung ermittelt werden (Kap. 2.3.1, Abb. 2.19), die der des Gesamtmusters entspricht. Alle drei Verfahren können die *barberplaid illusion* erklären, doch sie können nicht zur Erklärung der *barberpole illusion* herangezogen werden, da am Balken nur Bewegungsvektoren orthogonal zur Balkenorientierung existieren. Die Hypothese, daß Terminatoren eine wichtige Rolle bei der Bewegungsverarbeitung spielen, kann hingegen beide Illusionen erklären (Hildreth 1984, Nakayama und Silverman 1988b, Vallortigara und Bressan 1991). Das vorliegende Modell unterstützt diese Annahme. Es wird die Bewegungsrichtung der Terminatoren extrahiert und zur Beeinflussung der Bewegungsrichtungsdetektion an den Konturkanten hinzugezogen, so daß an allen Teilen des Musters die gleiche Bewegungsrichtung signalisiert wird.

Grossberg und Mingolla stellten (1993) ein Modell vor, welches Linienenden in die Bewegungsverarbeitung einband. Das in dieser Arbeit vorgestellte Modell unterscheidet sich von deren in Funktionsweise und Ergebnis. Grossberg und Mingolla bestimmten die Bewegungsrichtung der Linienenden nicht explizit, sondern als Teil einer Kontur, wodurch die Bewegungsverarbeitung der Linienenden immer noch - wenn auch nur gering - dem Korrespondenzproblem unterliegt. Die stärkste Bewegungsdetektorantwort an den Linienenden entspricht der der zugehörigen Kante, das ist in der Regel eine Bewegungsrichtung orthogonal zur Kantenorientierung. Aus der nachfolgenden räumlichen und zeitlichen Mittelwertbildung resultiert die Bewegungsrichtung der Linienenden. Diese von Grossberg und Mingolla recht weiträumig durchgeführte Mittelung bringt allerdings eine mangelhafte Genauigkeit bei der Lokalisation der Linienenden mit sich. Die Bewegungsrichtungsdetektion entlang der Kontur bleibt von der Detektion der Linienendenbewegungsrichtung unbeeinflusst. Dies ist natürlich nicht konform mit der Wahrnehmung, denn der Bewegungseindruck ist nicht nur an den Linienenden, sondern auch an den Konturkanten vorhanden. Außerdem ist die Lokalisation der Linienenden exakt möglich. Das in dieser Arbeit vorliegende Netzwerk mit impulskodierenden Modellneuronen extrahiert zunächst zusammenhängende Kanten und bestimmt deren Linienenden. Eine direkte Extraktion der Linienenden wäre im Vergleich dazu störanfälliger. Nachfolgend wird die Bewegungsrichtung der Linienenden bestimmt. Das Korrespondenzproblem ist bei einer sicheren Detektion der Linienenden am geringsten. Aus diesem Grunde wurde zusätzlich eine Rückkopplung von den Richtungsdetektoren für Linienenden zu den Linienendendetektoren implementiert. Hierbei kommt der modulatorisch wirkende Linking-

Eingang des Modellneurons ebenso zum Tragen, wie in der Beeinflussung der Richtungsdetektoren für Konturen. Die Richtungsdetektoren, deren Richtungspräferenz der Bewegungsrichtung der Linienenden entspricht, werden dadurch facilitiert, wodurch eine kohärente Bewegung entsteht.

4.4 Bewegungskонтast

In Kapitel 3.5 konnte gezeigt werden, daß im vorliegenden Modell eine Objektseparation aufgrund des Bewegungskontastes möglich ist. Eine Gruppierung von zum Beispiel sich bewegenden Flecken durch die Detektion der absoluten Bewegungsrichtung in getrennten neuronalen Schichten (Abb. 3.111) ist in vielen realen Fällen möglich, dennoch ist sie nur als Spezialfall anzusehen. In dieser Arbeit wurde ein allgemeingültiges Verfahren zur Objektseparation angewandt. Eine Segmentierung bewegter Szenen ist durch den lokalen Bewegungskontast möglich, wie zum Beispiel die Untersuchungen Nothdurfts (1993) und die in Abbildung 2.11 dargestellten gegenläufig rotierenden Kreisscheiben zeigen. Die in Kapitel 3.5 beschriebene Erweiterung des neuronalen Modells zur Bewegungsdetektion ermöglicht die Extraktion des lokalen Bewegungskontastes. Aktive Bewegungskontastdetektoren weisen hierbei auf Objektgrenzen hin. Eine mögliche Repräsentation durch synchrone neuronale Aktivität der einzelnen Objekte wurde aufbauend auf diesem Modell von Pauly et al. (1996) gezeigt. Die prinzipielle Verschaltungsstruktur der Kontrastextraktion im vorliegenden Modell läßt sich ebenso zur Texturanalyse einsetzen, wie von Weitzel et al. (1995) und Weitzel (1998) gezeigt wurde. Hierzu muß die Eingabe der Kontrastdetektoren geändert werden. Anstelle von Richtungsdetektoren müssen Orientierungsdetektoren den Kontrastdetektoren vorgeschaltet werden. Die Grenzen des in dieser Arbeit entwickelten Modells zur Bewegungskontastdetektion wurden in Abbildung 3.115 E deutlich. Sich zufällig bewegende Flecken führen zu einem hohen Bewegungskontast, der über das gesamte Objekt verteilt ist, aber zu keiner Objektgrenze. Für eine genauere Analyse der Objektkontur müßte das vorgestellte Modell noch um weitere kontrastdetektierende Schichten ergänzt werden. Eine Möglichkeit zur Kontrastbestimmung des Bewegungskontastes wurde von Fenske (1995) gezeigt. Eine andere Möglichkeit besteht in der Zuführung der Bewegungskontastinformation in die Orientierungs- und Linienendendetektion, denn schließlich kann auch eine auf Bewegungskontast beruhende Form als solche erkannt und klassifiziert werden. Hierbei kann ebenfalls angenommen werden, daß stationäre Bewegungskontaste zu einer oszillatorischen, synchronen Aktivität der Orientierungsdetektoren führen und eine Objektseparation im Sinne der Korrelationshypothese durchgeführt werden kann. Für die Texturanalyse wurde dies bereits von Weitzel (1998) gezeigt.

Im vorliegenden Modell wird der Bewegungskontast auf zwei unterschiedliche Arten extrahiert. Zum einen durch ein inhibitorisch wirkendes Umfeld, in das Modellneuronen mit gleicher und ähnlicher Richtungspräferenz eingebunden sind, zum anderen durch ein facilitatorisch wirkendes Umfeld, in dem Modellneuronen mit stark unterschiedlicher Richtungspräferenz involviert sind. Neurophysiologische Untersuchungen deuten ebenso auf zwei Typen von MT-Neuronen zur lokalen Bewegungskontastdetektion hin. Ein Großteil der von Allman et al. (1985) untersuchten MT-Neuronen besaßen zum Beispiel ein rezeptives Feld mit Zentrum/Umfeld-Antagonismus, so daß die Aktivität dieser Neuronen sank, wenn sich in ihrer Umgebung Objekte in die gleiche Richtung bewegten wie im Zentrum. Diese Neuronen könnten somit genauso für die Extraktion des lokalen Bewegungskontastes verantwortlich sein, wie die MT-Neuronen, die von Tanaka et al. (1986) untersucht wurden. Diese zeigten eine Facilitation in ihrer Aktivität bei Reizung ihres Umfeldes mit Objekten, die sich entge-

gengesetzt zu der Bewegungsrichtung des Objektes im Zentrum bewegt. Beide vorgestellten Verfahren zur Bewegungskontrastdetektion sind somit biologisch plausibel.

4.5 Einsetzbarkeit in der technischen Bildverarbeitung

Das hier vorgestellte neuronale Netzwerk besitzt einen hohen Grad an Flexibilität. Es ist an keine bestimmte Aufgabe gebunden, sondern kann vielfältig eingesetzt werden. Das zum Beispiel in Graefe (1995) vorgestellte technische Bildverarbeitungssystem zur Fahrerunterstützung in Kraftfahrzeugen detektiert Verkehrsschilder, indem im rechten oberen Bereich des Eingabebildes Korrelationsberechnungen durchgeführt werden. Dieses System ist somit auf Schilder restringiert, die am rechten Straßenrand stehen. Bei Überholvorgängen befinden sich die so gesuchten Schilder nicht mehr im Verarbeitungsbereich und können somit nicht gefunden werden. Verkehrsschilder, die über der Fahrbahn angebracht sind, bleiben ebenso unberücksichtigt. Überholende Kraftfahrzeuge werden von Graefe durch vertikale Kanten, die links unten im Bildbereich auftauchen, identifiziert. Rechts überholende Kraftfahrzeuge werden genauso wenig berücksichtigt, wie zum Beispiel Fußgänger, die sich auf die Fahrbahn begeben. In Graefes System wird versucht, typische Situationen des Straßenverkehrs mit ganz speziellen Methoden zu bewältigen.

Das in dieser Arbeit vorgestellte Gesamtsystem basiert im Gegensatz zu Graefes auf einem allgemeinen Ansatz. Zur Überprüfung der Funktionsfähigkeit wurden wie bei Graefe typische Szenen des Straßenverkehrs herangezogen. Hierfür dienten zahlreiche Szenen mit Verkehrsschildern, Fußgängern und überholenden Kraftfahrzeugen. In den meisten Szenen wurde dabei die Kamera selbst bewegt. Es konnte in Kapitel 3.2.2 gezeigt werden, daß das Transientensystem die Bewegungsrichtung dieser Objekte für eine ausreichend lange Zeit bestimmen kann. Kann aus Gründen der Modellneuronenanzahl nur eine einzige Auflösung zur Bewegungskontrastdetektion benutzt werden, können zum Beispiel Verkehrsschilder nur detektiert werden, wenn sich das Kraftfahrzeug, in dem sich die Kamera befindet, mit einer Geschwindigkeit innerhalb eines bestimmten Intervalles bewegt. Bewegt sich das Kraftfahrzeug zum Beispiel zu langsam, so bewegt sich auch das Schild nur mit einer geringen Relativgeschwindigkeit, die als Projektion in das zweidimensionale Kamerabild außerhalb des für die Richtungsdetektoren einer bestimmten Auflösung charakteristischen Geschwindigkeitsbereiches liegt. Der Geschwindigkeitsbereich des Richtungsdetektors könnte aber dynamisch angepaßt werden, wenn die Anzahl der Interpolationsschritte zwischen zwei Einzelbildern entsprechend der momentanen Geschwindigkeit des Kraftfahrzeuges reduziert würde. Durch diese dynamische Anpassung könnten die Richtungsdetektoren, obwohl sie nur sensitiv für einen bestimmten projizierten Geschwindigkeitsbereich sind, einen großen Bereich an auftretenden Geschwindigkeiten abdecken.

Technische Bildverarbeitungssysteme zur Fahrerunterstützung müssen mobil einsetzbar sein und der Echtzeitanforderung gerecht werden. Da konventionelle Computer dieser Anforderung in der Regel nicht genügen, wurde in Zusammenarbeit mit der Gesamthochschule Paderborn ein Hardwaresystem zur beschleunigten Berechnung impuls-kodierender neuronaler Netze (Accelerator) entwickelt. Dieser Accelerator unterliegt gegenüber konventionellen Computern einigen Einschränkungen. Alle in dieser Arbeit aufgeführten Simulationen - dies gilt sowohl für technisch orientierte Simulationen als auch für die Simulationen der psychophysischen Effekte - wurden konform zu diesen Einschränkungen durchgeführt. Das Accelerator-konzept setzt eine geringe Aktivität voraus. In allen Simulationen blieb die Aktivität unter einem Prozent. Die größte Beschränkung liegt in der Anzahl der Modellneuronen. Aufgrund dieser Beschränkung konnte das Transientensystem nie vollständig eingesetzt werden. Auch eine Kopplung mit dem Kontur/Form-System, die für eine vollständige Objektdefinition nötig wäre, konnte aus diesem Grund nicht durchgeführt werden. Als Minimalanforderung an

ein technisches Gesamtsystem sind mindestens 10^6 Modellneuronen anzusetzen. Die Anzahl der zur Verfügung stehenden Synapsen war in den Simulationen ebenso ausreichend wie die Anzahl der möglichen Abklingkonstanten pro Teilpotential. Wünschenswert wäre eine größere Amplitudenauflösung der Teilpotentiale, damit im Hinblick auf das kleinste Dekrement auch große Zeitkonstanten bei der zeitlichen Tiefpaßfilterung möglich sind.

Auf die Implementierung von Spontanaktivität wurde aus zwei Gründen verzichtet. Zum einen kann die spontane Aktivität wegen der relativ kleinen Zeitkonstanten nicht kumuliert werden, zum anderen muß die Gesamtaktivität aus Gründen der Rechenzeit so gering wie möglich gehalten werden. Eine Spontanaktivität würde zudem zu einer Erregung vieler Teilpotentiale führen, die ohne diese nicht in die Abklingphase hätten miteinbezogen werden müssen. Das für die Objektseparation - das ist hier die zeitliche Trennung der neuronalen Aktivitäten - wichtige Rauschen (z.B. Stoecker 1993, Schott 1994) ist bereits in den Eingabebildern enthalten und muß nicht durch eine spontane Aktivierung erzeugt werden

4.6 Ausblick

Die Kopplung des Transientensystems mit dem Kontur/Form-System führt wegen des allgemeinen und an der Biologie orientierten Ansatzes zu einem flexiblen, universellen System zur Bildverarbeitung, das nicht nur technisch einsetzbar ist, sondern auch als Modell zur Erklärung psychophysischer und neurophysiologischer Untersuchungen herangezogen werden kann. Im biologischen Sehsystem spielt bereits vorhandenes Wissen über Objekte eine erhebliche Rolle. Deswegen sollte das im Rahmen des BMBF-Projektes „Elektronisches Auge“ entstandene System um weitere Module zur Objekterkennung und Invariantenbildung ergänzt werden. Diese Module müßten in der Lage sein, eine Invariantenbildung bezüglich Ort, Größe, Rotation und Form durchzuführen, um aus dieser mit Hilfe von Assoziativspeicher und Wissensbasis Objekte erkennen zu können. Zur Speicherung neuer Objekte und zur Anpassung an die Umwelt sollte das neuronale Netzwerk lernfähig sein. Mit diesen zusätzlichen Modulen zur Objekterkennung und einer geeigneten Lernregel hat dieses System sehr gute Chancen, gegen klassische Bildverarbeitungssysteme zu bestehen.

5 Literatur

Adelson M., Movshon J.A. (1982), *Phenomenal Coherence of Moving Visual Patterns*. Nature 300, pp. 523-525.

Adelson E.H., Bergen J.R. (1985), *Spatiotemporal Energy Models for the Perception of Motion*. J. Opt. Soc. Am. 2, pp. 284-299.

Alais D., Wenderoth P., Burke D. (1994), *The Contribution of One-Dimensional Motion Mechanisms to the Perceived Direction of Drifting Plaids and their Aftereffects*. Vision Res. 34 (14), pp. 1823-1834.

Alais D., Burke D., Wenderoth P. (1996a), *Further Evidence for Monocular Determinants of Perceived Plaid Direction*. Vision Res. 36 (9), pp. 1247-1253.

Alais D., van der Smagt M.J., Verstraten F.A.J., van de Grind W.A. (1996b), *Monocular Mechanisms Determine Plaid Motion Coherence*. Visual Neuroscience 13, pp. 615-626.

Albright T.D. (1984), *Direction and Orientation Selectivity of Neurons in visual Area MT of the Macaque*. J. Neurophysiol. 52, pp. 1106-1130.

Albright T.D., Desimone R., Gross C.G. (1984), *Columnar Organization of Directionally Selective Cells in Visual Area MT of the Macaque*. J. Neurophysiol. 51, pp. 16-31.

Allman J.M., Kaas J.H. (1971), *A Representation of the Visual Field in the Caudal Third of the Middle Temporal Gyrus of the Owl Monkey (Aotus Trivirgatus)*. Brain Res. 31, pp. 85-105.

Allman J.M., Miezin F., McGuinness E. (1985), *Direction- and Velocity-Specific Responses from Beyond the Classical Receptive Field in the Middle Temporal Visual Area (MT)*. Perception 14, pp. 105-126.

Anstis S. (1986), *Phi-Movement as a Subtraction Process*. Vision Res. 10, pp. 1411-1430.

Anstis S., Ramachandran V.S. (1987), *Visual Inertia in Apparent Motion*. Vision Res. 27 (5), pp. 755-764.

Arndt M. (1993), *Repräsentation räumlicher und zeitlicher Stetigkeit durch Synchronisation neuronaler Signale*. Dissertation, Universität Marburg, FB Physik.

Barlow H.B., Hill R.M. (1963), *Selective Sensitivity to Direction of Motion in Ganglion Cells in the Rabbit's Retina*. Science 139, pp. 412-414.

Barlow H.B., Levick W.R. (1965), *The Mechanism of Directionally Selective Units in the Rabbit's Retina*. J. Physiol. 178, pp. 477-504.

Barlow H.B., Hill R.M., Levick W.R. (1964), *Retinal Ganglion Cells Responding Selectively to Direction and Speed of Image Motion in the Rabbit*. J. Physiol. 173, pp. 377-407.

- Becker K. (1995)**, *Computersimulation der Vertebratenretina mit biologienahen Neuronen*. Unveröffentlichte Diplomarbeit, Universität Marburg, FB Physik.
- Behring J. (1998)**, *Kodierung visueller Objekte durch synchronisierte Aktivität lokaler Merkmalsdetektoren*. Unveröffentlichte Diplomarbeit, Universität Marburg, FB Physik.
- Beutter B.R., Mulligan J.B., Stone L.S. (1996)**, *The Barberplaid Illusion: Plaid Motion is Biased by Elongated Apertures*. Vision Res. 36 (19), pp. 3061-3075.
- Biederman I. (1987)**, *Recognition-by-Components: A Theory of Human Image Understanding*. Psych. Review Vol. 94, pp. 115-147.
- Bowling D.B. (1994)**, *Timing Differences Between the Light Responses of X Cells Recorded Simultaneously in Cat Lateral Geniculate Nucleus*. Visual Neuroscience 2, pp. 383-389.
- Boycott B.B., Wässle H. (1974)**, *The Morphological Types of Ganglion Cells of the Domestic Cat's Retina*. J. Physiol. 240, pp. 397-419.
- Braddick O. (1974)**, *A Short-Range Process in Apparent Motion*, Vision Res. Vol. 14, pp. 519-527.
- Braddick O. (1993)**, *Segmentation Versus Integration in Visual Motion Processing*. TINS 16 (7), pp. 263-268.
- Bradley D.C., Qian N., Andersen R.A. (1995)**, *Integration of Motion and Stereopsis in Cortical Area MT of the Macaque*. Nature 373, pp. 609-611.
- Czambor S. (1997)**, *Adaptation raum-zeitlicher Übertragungseigenschaften in Netzwerken aus impulskodierenden Neuronen*. Unveröffentlichte Diplomarbeit, Universität Marburg, FB Physik.
- Chang J.J., Julesz B. (1984)**, Vision Res. 24, pp. 1781-1788.
- Chey J., Grossberg S., Mingolla E. (1998)**, *Neural Dynamics of Motion Processing and Speed Discrimination*. Vision Res. 38, pp. 2769-2786.
- Cleland B.G., Levick W.R., Wässle H. (1975)**, *Physiological Identification of a Morphological Class of Cat Retinal Ganglion Cells*. J. Physiol. 248, p.151.
- Dacey D.M., Petersen M.R. (1992)**, *Dendritic Field Size and Morphology of Midget and Parasol Ganglion Cells of the Human Retina*. Proc. Natl. Acad. Sci. 89, pp. 9666-9670.
- Dicke P.W. (1992)**, *Simulation dynamischer Merkmalskopplungen in einem neuronalen Netzwerkmodell*. Dissertation, Universität Marburg, FB Physik.
- Dinse H.R., Krüger K. (1994)**, *The Timing of Processing Along the Visual Pathway in the Cat*. NeuroReport 5, pp. 893-897.
- Dubner R., Zeki S.M. (1971)**, *Response Properties and Receptive Fields of Cells in an Anatomically Defined Region of the Superior Temporal Sulcus in the Monkey*. Brain Res. 35, pp. 528-532.

- Duffy C.J., Wurtz R.H. (1991)**, *Sensitivity of MST Neurons to Optic Flow Stimuli. II. Mechanisms of Response Selectivity Revealed by Small-Field Stimuli*. J. Neurophysiol. 65 (6), pp. 1346-1359.
- Duffy C.J., Wurtz R.H. (1995)**, *Response of Monkey MST Neurons to Optic Flow Stimuli with Shifted Centers of Motion*. The Journal of Neuroscience 15 (7), pp. 5192-5208.
- Eckhorn R. (1994)**, Persönliche Kommunikation, Universität Marburg, FB Physik.
- Eckhorn R., Bauer R., Brosch M., Jordan W., Kruse W., Munk M., Reitboeck H.J. (1988)**, *Coherent Oscillations: A Mechanism of Feature Linking in the Visual Cortex?* Biol. Cybern. 60, pp. 121-130.
- Eckhorn R., Reitboeck H.J., Arndt M., Dicke P. (1990)**, *Feature Linking via Synchronization Among Distributed Assemblies: Simulations of Results from Cat Visual Cortex*. Neural Computation 2, pp. 293-307.
- Egelhaaf M., Reichardt W. (1987)**, *Dynamic Response Properties of Movement Detectors: Theoretical Analysis and Electrophysiological Investigation in the Visual System of the Fly*. Biol. Cyber. 56, pp. 69-87.
- Egelhaaf M., Hausen K., Reichardt W., Wehrhahn C. (1988)**, *Visual Course Control in Flies Relies on Neuronal Computation of Object and Background Motion*. TINS 11 (8), pp. 351-358.
- Enroth-Cugell C., Robson J.G. (1966)**, *The Contrast Sensitivity of Retinal Ganglion Cells of the Cat*. J. Physiol. 187, pp. 517-552.
- van Essen D.C., Gallant J.L. (1994)**, *Neural Mechanisms of Form and Motion Processing in the Primate Visual System*. Neuron 13, pp. 1-10.
- Famiglietti E., Kolb H. (1976)**, *Structural Basis for ON- and OFF-Center Responses in Retinal Ganglion Cells*. Science 194, pp. 193-195.
- Fenske J., Schott U., Stoecker M., Reitboeck H.J. (1995)**, *Feature Contrast I. A Neural Network for Pattern Segmentation via Motion Contrast Detection*. N.Elsner, R. Menzel (Ed.): Proceedings of the 23rd Göttingen Neurobiology Conference 1995, Volume II, Georg Thieme Verlag Stuttgart, New York, p. 889.
- von Foerster (1966)**, *Neural Networks*. Biological Computer Laboratory, Report No. 5.1, 1. April 1966, pp. 46.
- Frank G. (1997)**, *Ein digitales Hardwaresystem zur echtzeitfähigen Simulation biologienaher neuronaler Netze*. Dissertation, HNI-Verlagsschriftenreihe, Band 26, Grundlagen der Elektrotechnik, Herausgeber G.Hartmann.
- Frank G., Bilau N., Hartmann G. (1995)**, *Hardware Accelerator zur Simulation pulscodierter Neuronaler Netze*. In: G. Sagerer, S. Posch, F. Kummert (eds.): Mustererkennung 1995. Informatik aktuell. Springer Verlag, Berlin, pp. 194-201.

- Frank G., Hartmann G. (1995)**, *An Artificial Neural Network Accelerator for Pulse-Coded Model Neurons*. Proc. ICNN 95, Perth, Australien, Vol. 4, pp. 2014-2018.
- Francis G., Grossberg S. (1996)**, *Cortical Dynamics of Form and Motion Integration: Persistence, Apparent Motion, and Illusory Contours*. Vision Res. 36 (1), pp. 149-173.
- French A.S., Stein R.B. (1970)**, *A Flexible Neural Analog Using Integrated Circuits*. IEEE Vol. BME-17, No. 3, pp. 248-253.
- Frien A., Eckhorn R., Wölbern T., Bauer R. (1996)**, *Synchronization of Fast Oscillations in V1 of Awake Monkey Depends on Cortical Distance and on Angle Between Preferred Orientations*. In: Proceedings of the 24th Göttingen Neurobiology Conference 1996, Volume II, Georg Thieme Verlag Stuttgart, New York, p. 417.
- Fries W., Keizer K., Kuypers H.G.J. (1985)**, *Large Layer V1 Cells in Macaque Striate Cortex (Meynert Cells) Project to Both Superior Colliculus and Prestriate Visual Area V5*. Exp. Brain. Res. 58, pp. 613-616.
- Gabriel A. (1995)**, Persönliche Kommunikation, Universität Marburg, FB Physik.
- Gegenbach V., Nagel H. (1994)**, Persönliche Kommunikation, FhG-IITB Karlsruhe.
- Gengerelli J.A. (1948)**, *Apparent Movement in Relation to Homogenous and Heterogeneous Stimulations of the Cerebral Hemispheres*. J. Exp. Psychol. 38, pp. 592-599.
- Glünder H., Lehmann A. (1992)**, *Relaxation in 4D State Space - A Competitive Network Approach to Object-Related Velocity-Field Correction*. Artificial Neural Networks 2, pp. 509-512.
- Graefe V. (1995)**, *Merkmalsextraktion, Objekterkennung, Echtzeit-Bildverarbeitungssysteme*. In: Sichtsystemgestützte Fahrzeugführung und Fahrer-Fahrzeug-Wechselwirkung, Vol. I., H.H. Nagel (eds.), Infix-Verlag, pp. 121-192.
- Gray C. M., Koenig P., Engel A. K., Singer W (1989)**, *Oscillatory Responses in Cat Visual Cortex Exhibit Inter-Columnar Synchronization Which Reflects Global Stimulus Properties*. Nature, pp. 334-337.
- Green M. (1986)**, *What Determines Correspondence Strength in Apparent Motion?* Vision Res. 26 (4), pp. 599-607.
- Groh J.M., Born R.T., Newsome W.T. (1997)**, *How Is a Sensory Map Read Out? Effects of Microstimulation in Visual Area MT on Saccades and Smooth Pursuit Eye Movements*. The Journal of Neuroscience 17 (11), pp. 4312-4330.
- Grossberg S., Mingolla E.(1993)**, *Neural Dynamics of Motion Perception: Direction Fields, Apertures, and Resonant Grouping*. Perception & Psychophysics 53 (3), pp. 243-278.
- Grossberg S., Rudd M.E. (1992)**, *Cortical Dynamics of Visual Motion Perception: Short-Range and Long-Range Apparent Motion*. Psychological Review 99 (1), pp. 78-121.

Grzywacz N.M., Yuille A.L. (1991), *Theories for the Visual Perception of Local Velocity and Coherent Motion*. In: Computational Models of Visual Processing, M.S. Landy, J.A. Movshon (eds.), MIT Press, pp. 231-252.

Grzywacz N.M., Watamaniuk S.N.J., McKee S.P. (1995), *Temporal Coherence Theory for the Detection and Measurement of Visual Motion*. Vision Res. 35 (22), pp. 3183-3203.

Hammond P., MacKay D.M. (1981), *Modulatory Influences of Moving Textured Backgrounds on Responsiveness of Simple Cells in Feline Striate Cortex*. J. Physiol. 319, pp. 431-442.

Hammond P., Smith A.T. (1983), *Directional Tuning Interactions Between Moving Oriented and Textured Stimuli in Complex Cells of Feline Striate Cortex*. J. Physiol. 342, pp. 35-49.

Hartmann G. (1982), *Recursive Features of Circular Receptive Fields*. Biol. Cybern. 43, pp. 199-208.

Hawken M.J., Parker A.J., Lund J.S. (1988), *Laminar Organization and Contrast Sensitivity of Direction-Selective Cells in the Striate Cortex of the Old World Monkey*. J. Neurosci. 8, pp. 3541-3548

Hebb D.O. (1949), *The Organization of Behavior*. John Wiley, New York 1949.

Heeley D.W., Buchanan-Smith H.M. (1994), *Changes in the Perceived Direction of Drifting Plaids, Induced by Asymmetrical Changes in the Spatio-Temporal Structure of the Underlying Components*. Vision Res. 34 (6), pp. 775-797.

Heggelund P. (1981), *Receptive Field Organization of Simple Cells in Cat Striate Cortex*. Exp. Brain Res. 42, pp. 89-98.

Heitger F., Rosenthaler L., von der Heydt R., Peterhans E. (1992), *Simulation of Neural Mechanisms from Simple to End-Stopped Cells*. Vision Res. 32, pp. 963-981.

Hildreth E.C. (1984), *Computations Underlying the Measurement of Visual Motion*. Artificial Intelligence 23, pp. 309-354.

Hildreth E.C., Ando H., Andersen R.A., Treue S. (1995), *Recovering Three-Dimensional Structure from Motion with Surface Reconstruction*. Vision Res. 35 (1), pp. 117-137.

Hirsch J.A., Gilbert C.D. (1991), *Synaptic Physiology of Horizontal Connections in the Cat's Visual Cortex*. J. Neuroscience 11 (6), pp. 1800-1809.

Hodgkin A.L., Katz B. (1949), *The Effect of Sodium Ions on the Electrical Activity of the Giant Axon of the Squid*. J. Physiol. 108, pp. 37-77.

Hodgkin A.L., Huxley A.F. (1952), *A Quantitative Description of Membrane Current and its Application to Conduction and Excitation in Nerve*. J. Physiol. 117, pp. 500-544.

Hubel D., Wiesel T. (1960), *Receptive Fields of Optic Nerve Fibers in the Spider Monkey*. J. Physiol. 154, pp. 572-580.

Hubel D., Wiesel T. (1965), *Receptive Fields and Functional Architecture in Two Nonstriate Visual Areas (18 and 19) of the Cat*. J. Neurophysiol. 26, pp. 229-289.

Hubel D., Wiesel T. (1968), *Receptive Fields and Functional Architecture of the Monkey Striate Cortex*. J. Physiol. 195, pp. 215-243.

Hummel J.E., Biederman I. (1992), *Dynamic Binding in a Neural Network for Shape Recognition*. Psych. Review Vol. 99, pp. 480-517.

Husain M., Treue S., Andersen R.A. (1989), *Surface Interpolation in Three-Dimensional Structure from Motion Perception*. Neural Computation 1, pp. 324-333.

Ikeda H., Wright M.J. (1974), *Evidence for "Sustained" and "Transient" Neurones in the Cat's Visual Cortex*. Vision Res. 14, pp. 133-136.

Julesz B. (1960), *Binocular Depth Perception of Computergenerated Patterns*. Bell Syst. Tech. J. 39, pp. 1125-1162.

Julesz B. (1971), *Foundations of Cyclopean Perception*. University of Chicago Press.

Kämpf D. (1998), *Selbstorganisation zweistufiger neuronaler visueller Repräsentation*. Unveröffentlichte Diplomarbeit, Universität Marburg, FB Physik.

MacKay D.M. (1961), Nature 192, pp. 739-740.

Koffka K. (1935), *Principles of Gestalt Psychology*. Harcourt, Brace & World, New York.

Kooi F.L. (1993), *Local Direction of Edge Motion Causes and Abolishes the Barberpole Illusion*. Vision Res. 33 (16), pp. 2347-2351.

Kolb H. (1994), *The Architecture of Functional Neural Circuits in the Vertebrate Retina*. Invest. Ophthalm. & Visual. Science. 35, pp. 2385-2404.

Kolb H., Nelson R., Mariani A. (1981), *Amacrine Cells, Bipolar Cells and Ganglion Cells of the Cat Retina: A Golgi Study*. Vision Res. 21, pp. 1081-1114.

Kopecz K., Mohraz K., Eckhorn R. (1996), *Visualizing Information Processing in Neural Vision Networks by Backprojecting Detector Characteristics to Visual Space*. Technical Report June 1996, Universität Marburg, FB Physik.

Kuffler S.W. (1953), *Discharge Patterns and Functional Organization of Mammalian Retina*. J. Neurophysiol. 16, pp. 37-68

Kuffler S.W., Nicholls J.G., Martin A.R. (1984), *From Neuron to Brain*. Sinauer Associates Inc. Publ. Sunderland 1984, p. 10.

Kupper R. (1998), *Selbstorganisation in lokalen Populationen gekoppelter Neuronen*. Unveröffentlichte Diplomarbeit, Universität Marburg, FB Physik.

Lagae L., Raiguel S., Orban G.A. (1993), *Speed and Direction Selectivity of Macaque Middle Temporal Neurons*. J. Neurophysiol. 69 (1), pp. 19-39.

Lankheet M.J.M., Molenaar J., van de Grind W.A. (1989), *The Spike Generating Mechanism of Cat Retinal Ganglion Cells.* Vision Res. 29 (5), pp. 505-517.

Lappe M. (1996), *Functional Consequences of an Integration of Motion and Stereopsis in Area MT of Monkey Extrastriate Cortex.* Neural Computation 8, pp. 1449-1461.

Lindemann J. (1998), *Modellbasierte Objektverfolgung mit dynamischen neuronalen Netzwerken.* Unveröffentlichte Diplomarbeit, Universität Marburg, FB Physik.

Lorenceau J., Shiffrar M. (1992), *The Influence of Terminators on Motion Integration Across Space.* Vision Res. 32 (2), pp. 263-273.

Lüschow A., Nothdurft H.C. (1993), *Pop-Out of Orientation but no Pop-Out of Motion at Isoluminance.* Vision Res. 33 (1), pp. 91-104.

von der Malsburg C. (1981), *The Correlation Theory of Brain Function.* Int. Report 81-2, Dept. Of Neurobiology, MPI for Biophysical Chemistry, Göttingen.

von der Malsburg C., Buhmann J. (1992), *Sensory Segmentation with Coupled Neural Oscillators.* Biol. Cybern. 67, pp. 233-242.

Marr D., Ullman S. (1981), *Directional Selectivity and Its Use in Early Visual Processing.* Proceedings of the Royal Society of London Series B, pp. 151-180.

Maunsell J.H.R., van Essen D.C. (1983), *Functional Properties of Neurons in Middle Temporal Visual Area of the Macaque Monkey. I. Selectivity for Stimulus Direction, Speed, and Orientation.* J. Neurophysiol. 49, pp. 1127-1147.

Maunsell J.H.R., Newsome W.T. (1987), *Visual Processing in Monkey Extrastriate Cortex.* Annual Review of Neuroscience 10, pp. 363-401.

Mikami A. (1992), *Spatiotemporal Characteristics of Direction-Selective Neurons in the Middle Temporal Visual Area.* Exp. Brain Res., pp. 40-46.

Mikami A., Newsome W.T., Wurtz R.H. (1986), *Motion Selectivity in Macaque Visual Cortex. I. Mechanisms of Direction and Speed Selectivity in Extrastriate Area MT.* J. Neurophysiol. 55, No. 6, pp. 1308-1327.

Möller M. (1995), *Leistungsvergleich verschiedener Neuronenmodelle bei grundlegenden Aufgaben der visuellen Informationsverarbeitung.* Unveröffentlichte Diplomarbeit, Universität Marburg, FB Physik.

Movshon J.A., Thompson I.D., Tolhurst D.J. (1978), *Receptive Field Organization of Complex Cells in the Cat's Striate Cortex.* J. Physiol. 283, pp. 79-99.

Nagel H. (1985), *Analyse und Interpretation von Bildfolgen.* Informatik-Spektrum 8, pp. 178-200, pp. 312-327.

Nakamura K., Matsumoto K., Mikami A., Kubota K. (1994), *Visual Response Properties of Single Neurons in the Temporal Pole of Behaving Monkeys.* J. Neurophysiol. 71 (3), pp. 1206-1221.

- Nakayama K. (1985)**, *Biological Image Motion Processing: A Review*. Vision Res. 25 (5), pp. 625-660.
- Nakayama K., Silverman G.H. (1988a)**, *The Aperture Problem - I. Perception of Non-rigidity and Motion Direction in Translating Sinusoidal Lines*. Vision Res. 28 (6), pp. 739-746.
- Nakayama K., Silverman G.H. (1988b)**, *The Aperture Problem - II. Spatial Integration of Velocity Information Along Contours*. Vision Res. 28 (6), pp. 747-753.
- Nawrot M., Sekuler R. (1990)**, *Assimilation and Contrast in Motion Perception: Explorations in Cooperativity*. Vision Res. 30 (10), pp. 1439 -1451.
- Nawrot M., Shannon E., Rizzo M. (1996)**, *The Relative Efficacy of Cues for Two-Dimensional Shape Perception*. Vision Res. 36 (8), pp. 1141-1152.
- Nothdurft H.C. (1993)**, *The Role of Features in Preattentive Vision: Comparison of Orientation, Motion and Color Cues*. Vision Res. 33 (14), pp. 1937-1958.
- Nothdurft H.C., Parlitz D. (1993)**, *Absence of Express Saccades to Texture or Motion Defined Targets*. Vision Res. 33 (10), pp. 1367-1383.
- Ohtani Y., Takemoto A., Kaihara T., Ejima Y. (1993)**, *On the Loss of Apparent Motion Between Isolated Chromatic Stimuli near Isoluminance*. Vision Res. 33 (7), pp. 935-938.
- Ostkamp L. (1996)**, *Bildvorverarbeitung realer Szenen mit dynamischen Modellneuronen zur kontextabhängigen Merkmalsextraktion, Merkmalskopplung und Szenensegmentierung*. Dissertation, Universität Marburg, FB Physik.
- Pauly M., Schott U., Eckhorn R., Reitboeck H.J. (1996)**, *Integration of Higher Order Feature Contrast Information for Dynamic Scene Segmentation with Spiking Neurons*. In: Proceedings of the 24th Göttingen Neurobiology Conference, Vol. II, N. Elsner and H. Schnitzler (eds.), Thieme Verlag, Stuttgart, New York, p. 794.
- Pauly M., Kopecz K., Eckhorn R. (1997)**, *A Dynamic Gaze Control Model Network of Pulse-Coding Neurons: Concurrent Treatment of Saccades and Smooth Pursuit*. In: Proceedings of the 25th Göttingen Neurobiology Conference, Vol. II, N. Elsner and R. Wässle (eds.), Thieme Verlag, Stuttgart, New York, p. 1022.
- Pauly M., Kopecz K., Eckhorn R. (1998)**, *Lateral Coupling Preserves Object Contours by Reducing 'Natural' Types of Noise in Models of Orientation Detector Maps*. In: Proceedings of the 26th Göttingen Neurobiology Conference, Vol. II, N. Elsner and R. Wehner (eds.), Thieme Verlag, Stuttgart, New York, p. 769.
- Peterhans E., von der Heydt R., Baumgartner G. (1986)**, *Mechanisms of Contour Perception in Monkey Visual Cortex. II. Contours Bridging Gaps*. J. Neuroscience 9 (5), pp. 1749-1763.
- Pinkus A., Pantle A. (1997)**, *Probing Visual Motion Signals with a Priming Paradigm*. Vision Res. 37 (5), pp. 541-552.

Poggio T., Reichardt W. (1973), *Consideration on Models of Movement Detection*. Kybernetik 13, pp. 223-227.

Qian N., Andersen R.A., Adelson E.H. (1994), *Transparent Motion Perception as Detection of Unbalanced Motion Signals I: Psychophysics*. J. Neuroscience 14, pp. 7357-7366.

Ramachandran V.S., Anstis S.M. (1983), *Extrapolation of Motion Path in Human Visual Perception*. Vision Res. 23, pp. 83-85.

Ramachandran V.S., Cavanagh P. (1987), *Motion Capture Anisotropy*. Vision Res. 27, pp. 97-106.

Rausch M., Erb M., Arndt M., Reitboeck H.J. (1993), *A Dynamic Associative Memory: Synchronization and Segmentation for Pattern Discrimination and Completion*. In: Gene-Brain-Behaviour, N. Elsner, M. Heisenberg (eds.), Thieme Verlag, Stuttgart, New York, p.872.

Reichardt W. (1957), *Autokorrelationsauswertung als Funktionsprinzip des Zentralnervensystems bei der optischen Wahrnehmung eines Insektes*. Z. Naturforschung B12, pp. 448-457.

Reichardt W. (1961), *Autocorrelation, a Principle for the Evaluation of Sensory Information by the Central Nervous System*. In Sensory Communication (Edited by Rosenblith W.A.), Wiley, New York.

Reitboeck H.J. (1980), In: L. Altmann (1982): *Psychophysische Untersuchungen zur Rolle der Synchronisation bei der Mustererkennung mit Hilfe einer mikroprozessorgesteuerten LED-Matrix*. Unveröffentlichte Diplomarbeit, Universität Marburg, FB Physik.

Reitboeck H.J. (1983), *A Multi-Electrode Matrix for Studies of Temporal Signal Correlations within Neural Assemblies*. In: Synergetics of the Brain, E.Basar, H.Flohr, H.Haken, and A.Mandell (eds.), Springer Verlag Berlin, pp. 1174-1182.

Reitboeck H.J., Stoecker M. (1993), *A Model for Scene Segmentation and Invariant Object Recognition in the Visual System*. 11th International Biophysics Conference, Budapest, Hungary.

Reitboeck H.J., Stoecker M., Hahn C. (1993), *Object Separation in Dynamic Neural Networks*. IEEE ICNN, San Francisco, USA, Vol. II, pp. 638-641.

Rodieck R.W. (1965), *Quantitative Analysis of Cat Retinal Ganglion Cell Response to Visual Stimuli*. Vision Res. 5, pp. 583-601.

Rodman H.R., Albright T.D. (1989), *Single-Unit Analysis of Pattern-Motion Selective Properties in the Middle Temporal Visual Area (MT)*. Exp. Brain. Res. 75, pp. 53-64.

Saam M., Schott U., Eckhorn R. (1997), *Extracting Relative Depth Information from Luminance Contrast and Motion Parallax by a Model Network of Spiking Neurons*. In: Proceedings of the 25th Göttingen Neurobiology Conference 1997, Vol. II, N. Elsner (eds.), Thieme Verlag, Stuttgart, New York, p. 1023.

Saam M., Eckhorn R. (1998), *Width of Association Fields in Visual Cortex is Influenced During Hebbian Learning by Signal Correlation Among Inputs and Lateral Spike Velocity: A Network Model with Spiking Neurons*. In: H. Lachnit, A. Jacobs, and F. Roesler (eds.), „Experimentelle Psychologie“, Pabst Science Publishers 1998, p. 298.

Salzman C.D., Britten K.H., Newsome W.T. (1990), *Cortical Microstimulation Influences Perceptual Judgements of Motion Direction*. *Nature* (346), pp. 174-177.

van Santen J.P.H., Sperling G. (1985), *Elaborated Reichardt Detectors*. *J. Opt. Soc. Am.* 2, pp. 300-321.

Schillen T.B., König P. (1991), *Stimulus-Dependent Assembly Formation of Oscillatory Responses: II. Desynchronisation*. *Neural Computation* 3, pp. 167-178.

Schmidt R.F. (1985), *Grundriß der Sinnesphysiologie*. Springer, Berlin, 1985.

Schott U. (1994), *Simultanes, unüberwachtes Lernen phasenseparierter Muster im rückgekoppelten Assoziativspeicher*. Unveröffentlichte Diplomarbeit, Universität Marburg, FB Physik.

Schott U., Eckhorn R., Reitboeck H.J. (1996), *Extracting Motion Contrast by a Neural Network of Spiking Neurons*. In: *Proceedings of the 24th Göttingen Neurobiology Conference 1996*, Vol. II, N. Elsner, H. Schnitzler (eds.), Thieme Verlag, Stuttgart, New York, p. 795.

Schott U., Eckhorn R., Reitboeck H.J. (1997), *Motion Capture by Contour Terminators via Feature Linking. A Simulation of the Barberpole Illusion with a Neural Network of Spiking Neurons*. In: *Proceedings of the 25th Göttingen Neurobiology Conference 1997*, Vol. II, N. Elsner (eds.), Thieme Verlag, Stuttgart, New York, p. 1025.

Schott U., Eckhorn R. (1998), *Contour Motion-Capture by Sharp Bends in a Network Simulation with Spiking Neurons*. *Beiträge zur 1. Tübinger Wahrnehmungskonferenz 1998*, Knirsch Verlag, p. 43.

Shechter S., Hochstein S. (1990), *ON and OFF Pathway Contributions to Apparent Motion Perception*. *Vision Res.* 30 (8), pp. 1189-1204.

Shimojo S., Silverman G.H., Nakayama K. (1989), *Occlusion and the Solution to the Aperture Problem for Motion*. *Vision Res.* 29, pp. 619-626.

Shipley T.F., Kellman P.J. (1997), *Spatio-Temporal Boundary Formation: the Role of Local Motion Signals in Boundary Perception*. *Vision Res.* 37 (10), pp. 1281-1293.

Siminoff R. (1980), *Modelling the Vertebrate Visual System*. *J. theor. Biol.* 86, pp. 307-346, pp. 673-708.

Simpson W.A. (1990), *The Use of Different Features by the Matching Process in Short-Range Motion*. *Vision Res.* 30 (10), pp. 1421-1428.

Singer W., Gray C. M. (1995), *Visual Feature Integration and the Temporal Correlation Hypothesis*. *Annual Review of Neuroscience*, pp. 555-586.

Smith A.T., Snowden R.J., Milne A.B (1994), *Is Global Motion Really Based on Spatial Integration of Local Motion Signals?* Vision Res. 34 (18), pp. 2425-2430.

Snowden R.J., Braddick O.J. (1989), *The Combination of Motion Signals Over Time*. Vision Res. 29 (11), pp. 1621-1630.

Spengler C. (1995), Persönliche Kommunikation, Universität Marburg, FB Physik.

Spengler C. (1996), *Relationale Kodierung von Objektmerkmalen für die visuelle Mustererkennung in biologienahen neuronalen Netzen*. Unveröffentlichte Diplomarbeit, Universität Marburg, FB Physik.

Spengler C., Schott U., Eckhorn R., Reitboeck H.J. (1996), *A Model of Visual Cortical Orientation Tuning, Including Modulation by Mutual Synchronization Among Spiking Neurons of Iso-Orientation Preferences and Inhibition Among Different Orientations*. In: Proceedings of the 24th Göttingen Neurobiology Conference 1996, Volume II, Georg Thieme Verlag Stuttgart, New York, p. 792.

Sterling P., Freed M., Smith G. (1986), *Microcircuitry and Functional Architecture of the Cat Retina*. TINS.

Stoecker M. (1993), *Höhere Mechanismen der visuellen Informationsverarbeitung in neuronalen Netzen*. Dissertation, Universität Marburg, FB Physik.

Stoecker M., Reitboeck H.J. (1993), *Scene Segmentation and Invariant Object Recognition in a Dynamic Neural Network*. In: Gene-Brain-Behaviour, N. Elsner, M. Heisenberg (eds.), Thieme Verlag, Stuttgart, New York, p. 868.

Stone L.S., Watson A.B., Mulligan J.B. (1990), *Effect of Contrast on the Perceived Direction of a Moving Plaid*. Vision Res. 30 (7), pp. 1049-1067.

Stoner G.R., Albright T.D. (1992), *Neural Correlates of Perceptual Motion Coherence*. Nature 358, pp. 412-414.

Stoner G. R., Albright T. D. (1993), *Image Segmentation Cues in Motion Processing: Implications for Modularity in Vision*. Journal of Cognitive Neuroscience, pp. 129-149.

Tanaka K., Hikosaka K., Saito H., Yukie M., Fukada Y., Iwai E. (1986), *Analysis of Local and Wide-Field Movements in the Superior Temporal Visual Areas of the Macaque Monkey*. J. Neuroscience 6 (1), pp. 134-144.

Ternus J. (1926), Psychol. Forsch. 7, pp. 81-136.

Thiel A. (1998), *Selbstorganisation modulatorischer, lateraler Verbindungen zwischen Merkmalsdetektoren und ihre Rolle bei der Informationsverarbeitung*. Unveröffentlichte Diplomarbeit, Universität Marburg, FB Physik.

Treue S., Andersen R.A., Ando H., Hildreth E.C. (1995), *Structure-from-Motion*. Vision Res. 35 (1), pp. 139-148.

- Treue S., Andersen R.A. (1996)**, *Neural Responses to Velocity Gradients in Macaque Cortical Area MT*. Visual Neuroscience 13, pp. 797-804.
- Treue S., Husain M., Andersen R.A. (1991)**, *Human Perception of Structure-from-Motion*. Vision Res. 31 (1), pp 59-75.
- Trueswell J.C., Hayhoe M.H. (1993)**, *Surface Segmentation Mechanisms and Motion Perception*. Vision Res. 33 (3), pp. 313-328.
- Vallortigara G., Bressan P. (1991)**, *Occlusion and the Perception of Coherent Motion*. Vision Res. 31 (11), pp. 1967-1978.
- Wässle H., Boycott B. (1991)**, *Functional Architectures of the Mammalian Retina*. Physiological Reviews 71 (2), pp. 447-480.
- Wang Y., Wang L., Li B., Wang L., Diao Y. (1995)**, *How is direction selectivity organized in the extrastriate visual area PMLS of the cat?* NeuroReport 6, pp. 1969-1974.
- Watamaniuk S.N.J., McKee S.P., Grzywacz N.M. (1995)**, *Detecting a Trajectory Embedded in Random-Direction Motion Noise*. Vision Res. 35 (1), pp. 65-77.
- Watanabe M., Rodieck R.W. (1989)**, *Parasol and Midget Ganglion Cells of the Primate Retina*. The Journal of Comparative Neurology 289, pp. 434-454.
- Weitzel L., Stoecker M., Schott U., Reitboeck H.J. (1995)**, *Feature Contrast: II. A Neural Network for Texture Segmentation via Orientation Contrast Detection*. N.Elsner, R. Menzel (eds.): Proceedings of the 23rd Göttingen Neurobiology Conference 1995, Volume II, Georg Thieme Verlag Stuttgart, New York, p. 890.
- Weitzel L., Ostkamp R., Eckhorn R., Reitboeck H.J. (1996)**, *Definition of Complex Visual Object Boundaries by Synchronized Activities in a Linking Architecture of Spiking Neurons*. In: Proceedings of the 24th Göttingen Neurobiology Conference 1996, Volume II, Georg Thieme Verlag Stuttgart, New York, p. 793.
- Weitzel L., Kopecz K., Spengler C., Eckhorn R. (1997a)**, *Contour Based Figure-Ground Segregation for Real World Images Modelled by a Linking Architecture of Pulse-Coding Neurons*. In: Proceedings of the 25th Göttingen Neurobiology Conference 1997, Vol. II, N. Elsner (eds.), Thieme Verlag, Stuttgart, New York, p. 1027
- Weitzel L., Kopecz K., Spengler C., Eckhorn R., Reitboeck H.J. (1997b)**, *Contour Segmentation with Recurrent Neural Networks of Pulse-Coding Neurons*. In: G. Sommer, K. Daniilidis, J. Pauli: Computer Analysis of Images and Patterns, Springer Verlag, pp. 337-344.
- Wiesel T.N., Hubel D.H. (1966)**, *Spatial and Chromatic Interactions in the Lateral Geniculate Body of the Rhesus Monkey*. J. Neurophysiol. 29, pp. 1115-1156.
- Williams D.W., Philips G., Sekuler R. (1986)**, Nature 324, pp. 253-255.
- Williams D.W., Sekuler R. (1984)**, *Coherent Global Motion Percepts from Stochastic Local Motion*. Vision Res. 24 (1), pp. 55-62.

Wilson H.R., Mast R. (1993), *Illusory Motion of Texture Boundaries*. Vision Res. 33 (10), pp. 1437-1446.

Young R.A. (1987), *The Gaussian Derivative Model for Machine Vision: Visual Cortex Simulation*. J. Opt. Soc., GMR-5323.

Yuille A.L., Grzywacz N.M. (1988), *A Computational Theory for the Perception of Coherent Visual Motion*. Nature Vol. 333 (6168), pp. 71-74.

Zeki S.M. (1974), *Functional Organization of a Visual Area in the Posterior Bank of the Superior Temporal Sulcus of the Rhesus Monkey*. J. Physiol. 236, pp. 549-573.

Zeki S.M. (1978), *Uniformity and Diversity of Structure and Function in Rhesus Monkey Prestriate Visual Cortex*. J. Physiol. 277, pp. 273-290.

Danksagung

An erster Stelle möchte ich mich bei meinen „Doktorvätern“ Prof. Dr. Dr. Reitböck und Prof. Dr. Eckhorn für die Möglichkeit bedanken, in der Arbeitsgruppe Neurophysik eine Doktorarbeit als wissenschaftlicher Mitarbeiter anfertigen zu können.

Allen Mitgliedern der Arbeitsgruppe möchte ich für die sehr angenehme Arbeitsatmosphäre und die zahlreichen Diskussionen danken. Dies gilt besonders für meine ehemaligen Zimmergenossen Lothar Weitzel, Mike Stöcker, Rüdiger Ostkamp, Claus Spengler, Jens Behring und Michael Großbach.

Frau Thomas möchte ich großen Dank für die organisatorische Betreuung aussprechen.

Für die Hilfestellung bei Hard- und Softwareproblemen danke ich Martin Pauly und Mirko Saam.

Für die gute Zusammenarbeit mit den externen Projektpartnern möchte ich stellvertretend Gerald Frank (Gesamthochschule Paderborn) und Axel Jahnke (TU Berlin) danken.

Monika Buckler und Karen Zhuber-Okrog danke ich für das anstrengende Korrekturlesen dieser Arbeit.

Schließlich möchte ich noch allen Sportlern im Renthof 7 - stellvertretend seien hier Uwe Thomas, Eddy Wölbern, David Kämpf, Karim Mohraz und Hans-Jörg Brinksmeyer genannt - für den sportlichen Ausgleich zur Doktorarbeit danken.